Proyecto final: Plan de migración a la nube.

Problemática del caso expuesto.

La empresa Química Del Santo quiere construir una arquitectura de datos y servicios utilizando AWS para tal fin.

Situación actual:

- Cada sector cuenta con sus propias hojas de cálculo como base de datos y el criterio para organizar las métricas dependen de cada uno. Al tener una RDS los datos serán cargados y procesados todos en un mismo lugar, generando una estandarización de los procesos.
- El sector de finanzas lleva a cabo una vez al año un inventario donde reagrupa los inventarios particulares de cada sector y así, obtiene un inventario de toda la mercadería. Generando la base de datos en la RDS los inventarios van a estar actualizados frecuentemente permitiendo hacer un seguimiento detallado del stock. Además, al implementar RDS Multi-AZ permitirá tener un respaldo en caso de que fallara la infraestructura de RDS, permitiendo tener un continuo funcionamiento de los servicios.
- Además, el sector de finanzas cuenta con el sistema Tango Gestión por medio del cual lleva adelante la administración de la compra/venta de la empresa, haciéndola funcionar como base de datos. Nuevamente, una RDS permitirá tener una sección detallada sobre los datos de compra/venta y al estar integrada con los demás sectores permitirá hacer un mejor seguimiento de las transacciones. Por otro lado, al utilizar S3 la empresa va a tener un gran almacenamiento de datos y de fácil acceso, en vez de tenerlo de forma física y sin saber dónde está exactamente un archivo requerido. Además, S3 va a permitir generar un backup en caso de generarse algún fortuito respecto a los datos y archivos almacenados.
- Se encuentra el sector de ventas internas y externas, cada uno administrado por un individuo en particular. Por medio de la administración de las credenciales de usuario se podrá hacer un seguimiento detallado de cada uno de los usuarios de la empresa. Además, como se va a generar una base de datos común ya no se encontraran los datos aislados permitiendo que cualquier usuario que lo necesite pueda solicitar los permisos necesarios para utilizarlos.

Descripción de la empresa.

La empresa Química Del Santo se dedica a la creación de insumos relacionados al monitoreo de la esterilización hospitalaria, se encuentra radicada en la zona del partido de General San Martín en la provincia de Buenos Aires. La empresa lleva 50 años y se utilizan sistemas transaccionales para la gestión de los productos de la empresa y sus clientes.

Los sectores de la empresa utilizan hojas de cálculo para generar sus propias bases de datos.

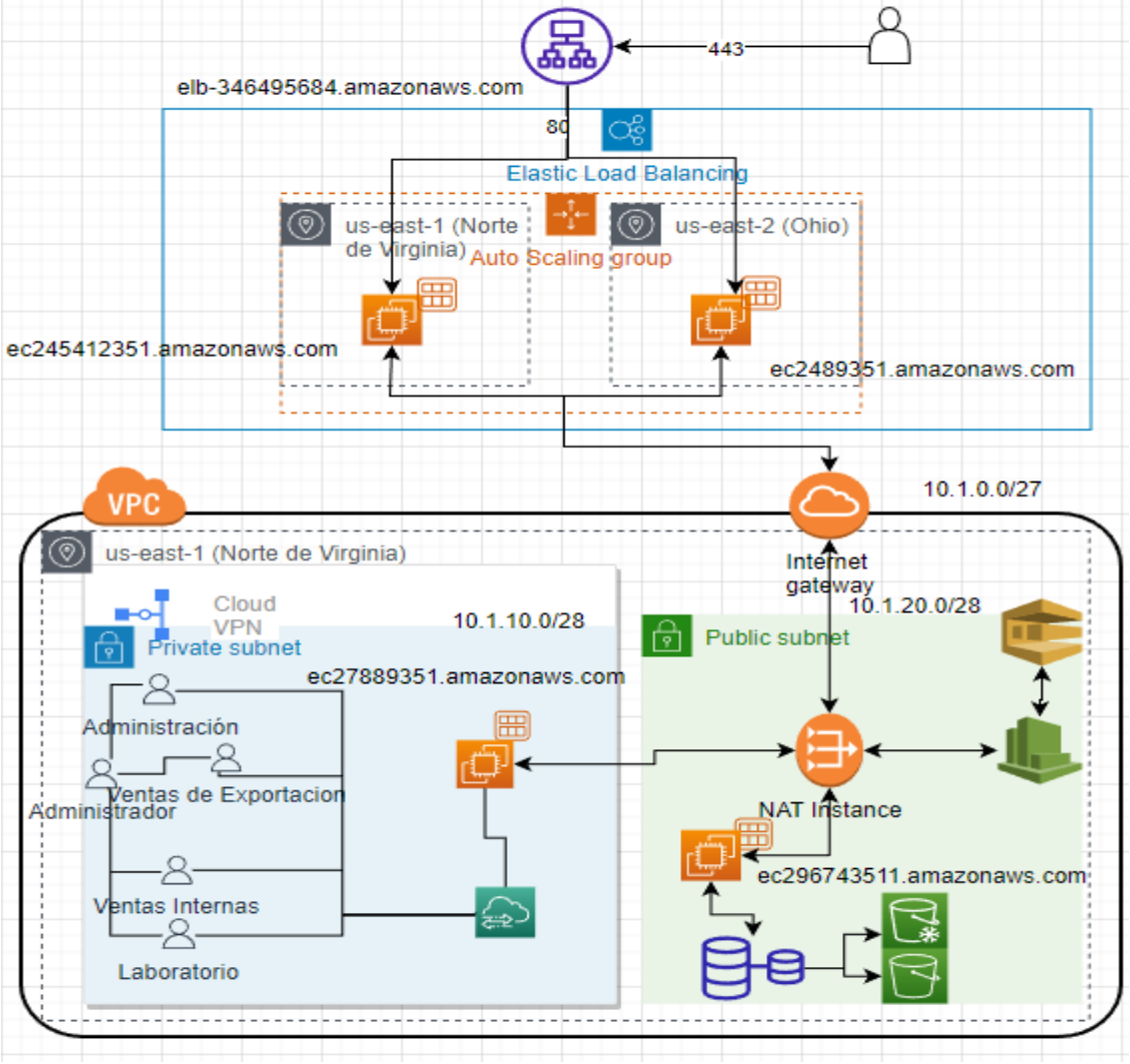
Dado el gran volumen de datos que la empresa manipula a diario y debido a que no hay un modelo establecido sobre como cargar los datos, cada sector lo lleva adelante a su manera.

La empresa busca obtener una mejor gestión, almacenamiento, procesamiento y administración de sus datos por medio del cómputo en la nube.

Objetivos de migración.

- Generar una estructura de IT dentro de la empresa donde cada miembro tenga un usuario con determinados permisos permitiendo realizar una trazabilidad.
- Tener una base de datos unificada donde cada sector pueda actualizarla diariamente si es necesario y donde los datos requeridos estén al alcance de todos los sectores.
- Modernizar la infraestructura de IT de la organización e incrementar el nivel de disponibilidad de los servicios críticos de la empresa.
- Generar un respaldo ante futuros acontecimientos que puedan generar un estado crítico en la empresa.
- Crear una arquitectura que permita el escalado y crecimiento de la compañía, desde la perspectiva de los servicios de IT.
- Pasar de una conexión de red privada de la empresa a una VPC, generando una subnet dividida entre los diferentes sectores de la empresa.

Arquitectura y descripción de los servicios de AWS



Arquitectura de servicios		
Servicio	Descripción	Justificación
IAM	AWS Identity and Access Management (IAM) es un servicio web que nos ayuda a controlar de forma segura el acceso a los recursos de AWS.	Resulta clave integrar dentro de nuestra arquitectura cloud el servicio de IAM, dado que nos permite controlar quién está autenticado (ha iniciado sesión) y si el usuario se encuentra autorizado (es decir si tiene permisos) para utilizar los recursos solicitados. Es crucial ya que se quieren generar diferentes tipos de usuarios con diferentes permisos dependiendo del sector, permitiendo un fácil seguimiento de los mismos.
EC2	Amazon Elastic Compute Cloud (Amazon EC2) proporciona capacidad de computación escalable en la nube de Amazon Web Services (AWS). Puede usarse para lanzar tantos servidores virtuales, configurar la seguridad y las redes, y administrar el almacenamiento.	El uso de Amazon EC2 elimina la necesidad de invertir inicialmente en hardware, de manera que puede desarrollar e implementar aplicaciones en menos tiempo y a un costo significativamente inferior. A su vez, puesto que la organización requiere el procesamiento de muchos datos, es factible integrar este servicio en clúster de máquinas virtuales. Además, se crearía un auto scaling group, lo que nos va a permitir crecer en cantidad de EC2 que se requieran en momentos críticos. También se crearía un launch template para agilizar a la hora de crear nuevas EC2 con AMI incorporada. Se generarán dos EC2 habilitadas en diferentes regiones de disponibilidad (con balanceadores de carga), una en us-east-1 (Norte de Virginia) y la otra en us-east-2 (Ohio) para el ingreso de los clientes en la plataforma de la empresa que son las que se encuentran con menor latencia respecto a la empresa (en caso de necesidad, si un cliente tiene demasiada latencia, se creará una instancia en otra zona de disponibilidad). Por otro lado, se crearán dos EC2 dentro de una VPC en la región de us-east-1 (Norte de Virginia) para el uso interno de la empresa. Todas las instancias de EC2 serán de t2.small.
DataSync	Es un servicio de transferencia de datos en línea que simplifica, automatiza y acelera la transferencia de datos entre los sistemas de almacenamiento en las instalaciones y los servicios de almacenamiento de AWS, así como entre los servicios de almacenamiento de AWS.	Se utilizara AWS DataSync para transferir los archivos onpremise de los diferentes usuarios de la empresa hacia la RDS por medio de la EC2
S3	Las clases de almacenamiento de Amazon S3 Glacier se crearon específicamente para el archivo de datos y están diseñadas con el objetivo de ofrecer el más alto rendimiento, la mayor flexibilidad de recuperación y el menor costo de almacenamiento de archivos en la nube posibles. S3 Estándar ofrece almacenamiento de objetos de alta durabilidad, disponibilidad y rendimiento para datos a los que se obtiene acceso con frecuencia.	Es necesario para el almacenamiento de archivos correspondientes a la compra y venta de los productos, como también de los registros de pruebas realizadas en el laboratorio ya que por normas que debe cumplir la empresa se tiene que tener un registro histórico de los mismos. Por lo tanto, un almacenamiento como Glacier donde no es necesario un continuo llamado de los archivos almacenados es perfecto. Por otro lado, tener un S3 estándar servirá para hacer un almacenamiento de los datos y archivos generados durante las semanas utilizando un versionado de S3 para evitar posibles inconvenientes. De esta manera se podría generar un snapshot por mes como backup y después del mes pasarlo a Glacier.

RDS	Amazon Relational Database Service (Amazon RDS) es un servicio web que facilita la configuración, la operación y la escala de una base de datos relacional en la nube de AWS. Proporciona una capacidad rentable y de tamaño ajustable para una base de datos relacional estándar y se ocupa de las tareas de administración de bases de datos comunes.	Se utilizará para generar una base de datos común a todos los usuarios ya que de esta manera cualquier información que se considere pertinente a todos los usuarios podría ser accesible sin tener que acceder a la base de datos física. Se utilizará la instancia de "rendimiento ampliable" por medio de MySQL para tal fin. Se utilizará una instancia de db.t2.large para tal fin y en caso de necesitarlo se cambiará a una mayor. Además, se utilizará RDS Multi-AZ. Así, en caso de que falle la infraestructura, la base de datos conmuta a una instancia en espera.
VPC	Amazon Virtual Private Cloud (Amazon VPC) le permite lanzar recursos de AWS en una red virtual que haya definido. Esta red virtual es muy similar a la red tradicional que usaría en su propio centro de datos, pero con los beneficios que supone utilizar la infraestructura escalable de AWS.	Se creará una VPC en la región us-east-1 (Norte de Virginia) que tendrá dos subredes, una privada donde estarán los usuarios de los diferentes sectores de la empresa que estarán conectados por medio de un EC2, y otra pública donde se creará un EC2 que le permitiría al EC2 de la subred privada pasar al sector público. El EC2 de la subred pública se encargaría de proporcionar los datos al RDS. Al mismo tiempo, se generará una NAT Instance que servirá de enlace entre la Internet Gateway y los sectores de EC2. Se creará un servicio de AWS Network Firewall para generar una mejor protección y monitoreo del tráfico de red, definiendo reglas de firewall y haciendo cumplir políticas para que se evite caer en dominios no autorizados.
CloudWatch	Monitorea los recursos y las aplicaciones de Amazon Web Services (AWS) que ejecuta en AWS en tiempo real. Puede utilizar CloudWatch para recopilar y hacer un seguimiento de métricas.	Se utilizará para ver las métricas de consumo tanto de billing como de CPU, RAM, disco, tráfico de red, los logs generados tanto en mi sistema como en la RDS permitiendo identificar quien hizo un determinado proceso. Pudiendo controlar todos estos consumos vamos a poder generar alertas que se disparen cuando algo exceda determinados parámetros o se generen consumos inusuales por determinados periodos establecidos haciendo que responda a estos eventos. Dentro de las EC2 se deberán instalar los agentes que permitan reportar si el sistema operativo está utilizando la memoria RAM y el disco que tienen reservado. Las métricas que se generarían son: Amazon API Gateway, AWS Backup, AWS Billing and Cost Management, AWS Client VPN, Amazon CloudWatch Logs, AWS DataSync, Amazon EC2, Amazon EC2 Auto Scaling, Elastic Load Balancing, Amazon Kinesis Data Analytics, Amazon Relational Database Service, Amazon Simple Queue Service, Simple Storage Service (Amazon S3), Amazon VPC.
SQS	Almacenar y recibir mensajes entre componentes de software de cualquier volumen, sin pérdida de mensajes ni la necesidad de que otros servicios estén disponibles.	Permitirá enviar mensajes de alerta que se disparen ante determinados tipos de eventos o alcance de métricas monitoreadas por CloudWatch.
Elastic Load Balancing	Distribuye automáticamente el tráfico entrante de la aplicación entre todas las instancias EC2 que están en ejecución. Elastic Load Balancing ayuda a administrar las solicitudes entrantes dirigiendo el tráfico de manera óptima para que ninguna instancia supere su capacidad.	Permitirá aumentar o disminuir la cantidad de EC2 activas, dependiendo de la cantidad requerida para momentos cruciales donde se generan picos de demanda, generando una mayor optimización de los servicios de manera horizontal.

Network Firewall	Conecta una política de firewall, que define el comportamiento de filtrado y supervisión del tráfico de red, a la VPC que desea proteger. La configuración del firewall incluye especificaciones para las zonas de disponibilidad y las subredes donde se ubican los extremos del firewall. También define configuraciones de alto nivel, como la configuración de registro del firewall y el etiquetado en el recurso de firewall de AWS.	Permitirá una mayor seguridad filtrando el tráfico de red de una política de firewall a la VPC. Además, se podrán implementar políticas y etiquetados particulares para las subredes permitiendo o denegando dependiendo del usuario y los accesos brindados.

Estructura de Billing

URL de calculadora para estimación de costos
Costo inicial: 0,00 USD
Costo mensual: 723,51 USD
Costo total de 12 meses: 8682,06 USD

Servicio	Costo mensual	Detalle de cálculo
EC2	76,91 USD	Ver Detalle
DataSync	1,25 USD	Ver Detalle
S3	60,09 USD	Ver Detalle
RDS	205,56 USD	Ver Detalle
VPC	235,47 USD	Ver Detalle
CloudWatch	40,60 USD	Ver Detalle
SQS	0,00 USD	Ver Detalle
Elastic Load Balancing	74,83 USD	Ver Detalle
Network Firewall	28,80 USD	Ver Detalle

Cronograma de implementación

	Mes	Mes 1				Mes 2			
	Semana	S1	S2	S3	S4	S1	S2	S3	S4
Actividad	Duración								
VPC	1 mes								
Network Firewall	1 mes								
IAM	1 semana								

SQS	1 semana								
EC2	3 semanas								
Elastic Load Balancing	3 semanas								
CloudWatch	2 semana								
DataSync	3 semanas								
RDS	1 mes								
S3	1 mes								

Ventajas e impacto en el negocio de la migración a Cloud

- Al contar con las actualizaciones y mantenimiento de los servicios por parte de AWS QDS se podría desligar de la contratación de personal para esos requerimientos.
- Se dejará de estar pendiente por si el hardware es suficiente o si se deberán hacer nuevas inversiones.
- Se pagará solo por lo que verdaderamente se utilice sin necesidad de comprar almacenamiento innecesario que al final no se utilizara.
- Se obtendrá una buena infraestructura de seguridad.
- Se obtendrá una base de datos con una estructura estandarizada para todos los usuarios
- Se generará un backup.
- Buen manejo de los usuarios, de los permisos y accesos permitiendo una excelente trazabilidad de los procesos.

Mejores prácticas

- Se generarán los usuarios por medio de script, utilizando terraform, logrando una automatización o mayor eficiencia a la hora de crearlos.
- El servidor se apagará durante la noche, salvo cuando se generen los backups que serán de forma mensual para pasarlo a S3 standart y anual para llevarlo a Glacier que se realizará por medio de un script que se correrá en terraform.
- Se realizará una buena política de permisos y tags dependiendo a qué sector de la empresa pertenezcan y en caso de ser necesario se otorgarán permisos dependiendo de la tarea solicitada. De esta forma se generaría un buen nivel de seguridad, administración y gestión en el ambiente de trabajo.
- Se utilizará Firewall y VPC para mejorar el sistema de seguridad en la nube. Además se establecerán subredes permitiendo un mejor aislamiento de los sistemas, generando que sea más difícil para alguien que logre entrar en el sistema, con propósitos maliciosos, poder atacar cada sector.
- Se realizará un seguimiento constante de las diferentes métricas por medio de CloudWatch, permitiendo una visualización detallada ante anomalías y configurar alertas de notificación, logrando una rápida acción ante situaciones críticas.

Conclusión

Por medio del plan de migración propuesto se logrará realizar una arquitectura mediante la cual la empresa mejorará en varios ámbitos. Ya sea el de seguridad por medio de la VPC, Network Firewall e IAM como en el caso de una mayor disponibilidad ante algún acontecimiento fortuito por medio del Multi-AZ de la RDS y Elastic Load Balancing. También, por medio de RDS, se obtendrá una base de datos implementada con MySQL donde se subirán los archivos por medio de DataSync para que luego de un procesamiento de los datos pasen a S3 donde serán almacenados de acuerdo a su antiguedad permitiendo realizar un backup de los mismos. Se logrará realizar un seguimiento de métricas para continuar mejorando a futuro en base al detalle de las herramientas implementadas y utilizadas gracias a CloudWatch y en caso de que alguna métrica sobrepase el nivel deseado de consumo se notificará por medio de SQS. Además, se beneficiaría por medio de EC2 que le permitirá contar con sistemas operativos actualizados ya que hoy en día no cuenta con ello y requeriría hacer una inversión tanto en hardware, software como de tiempo para su implementación. Por otro lado, solo pagaría por lo utilizado beneficiándose de no invertir en infraestructura que luego no usaría o preocupándose por si la infraestructura actual llegará a alcanzar.