

Lección 4

Amazon DynamoDB



Amazon DynamoDB

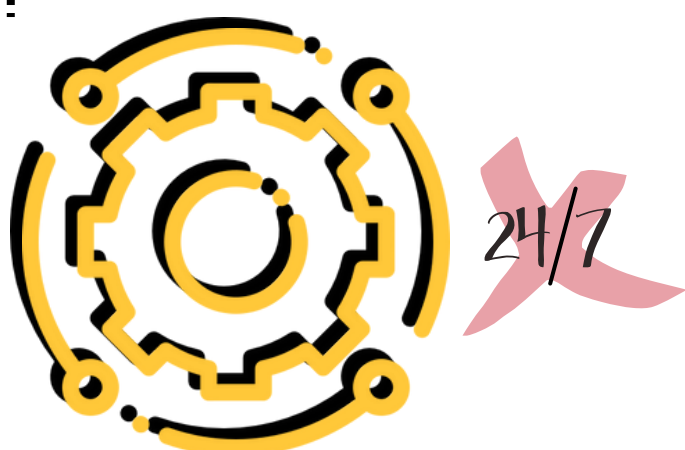


Amazon
DynamoDB

Un servicio de base de datos no relacional completamente administrada de clave-valor y documentos.



Eficacia a cualquier escala
Capacidad extrema de escalado horizontal



Sin servidor

Programación dirigida por eventos
(informática sin servidor)



Lista para el uso empresarial

Cifrado, controles de acceso,
copias de seguridad

Amazon DynamoDB es un servicio de base de datos NoSQL no relacional totalmente administrado por AWS. Destaca por ofrecer un rendimiento rápido y predecible, además de una escalabilidad sin complicaciones.

Algunos de los principales beneficios que ofrece son los siguientes:

- Escalabilidad masiva: Permite manejar una gran cantidad de tráfico y datos, manteniendo tiempos de respuesta consistentes en menos de diez milisegundos.
- Sin servidor: No requiere que se aprovisionen o administren servidores, ni que se instale o mantenga software alguno.
- Listo para empresas: Admite transacciones ACID, cifra todos los datos en reposo de manera predeterminada y ofrece funciones avanzadas como la replicación en varias regiones (tablas globales), control detallado de acceso e identidad, y copias de seguridad completas de los datos sin afectar el rendimiento.

Características de Amazon DynamoDB



Amazon
DynamoDB

Funciona bien para aplicaciones que cumplen con lo siguiente:

- Tienen grandes volúmenes de datos simples (rango alto de TB).
- Deben escalarse rápidamente.
- No necesitan combinaciones complejas.
- Requieren un rendimiento ultra alto y baja latencia.

Características clave

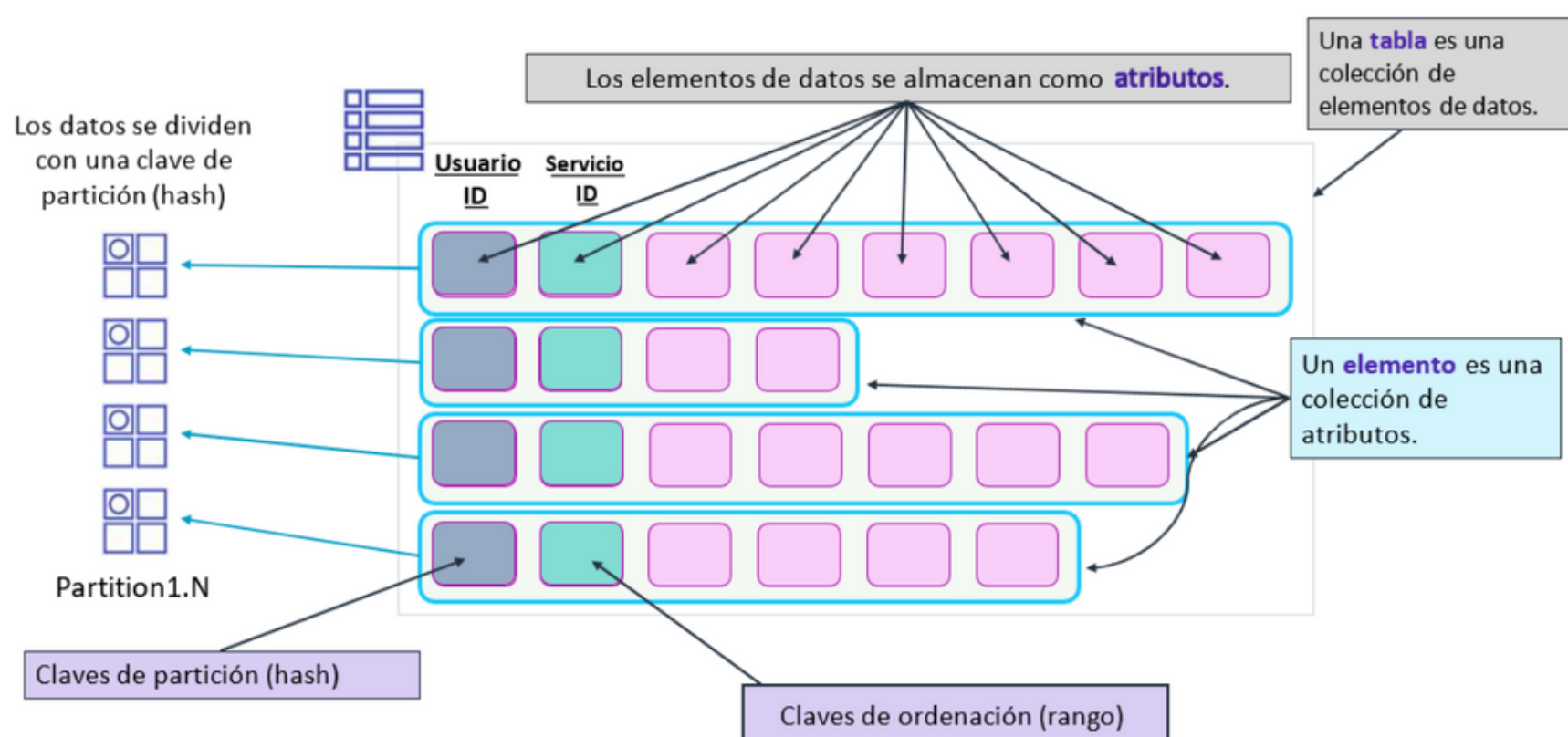
- Tablas NoSQL
- Los elementos pueden tener atributos diferentes
- Almacenamiento de caché en memoria
- Compatibilidad con picos de más de 20 millones de solicitudes por segundo

DynamoDB se destaca en entornos donde se necesita manejar grandes volúmenes de datos y escalarse rápidamente. Su rendimiento ultrarrápido y baja latencia lo hacen ideal para aplicaciones como videojuegos, tecnología publicitaria y aplicaciones móviles que requieren respuestas instantáneas. Sin embargo, si las cargas de trabajo implican operaciones complejas de consulta, es posible que una base de datos relacional sea más adecuada.

Una característica distintiva de DynamoDB es su flexibilidad en los esquemas de datos, ya que no tienen estructuras fijas. Esto significa que cada elemento puede contener un conjunto variable de atributos.

Debido a su capacidad para ofrecer almacenamiento en caché en memoria y manejar una cantidad extraordinaria de solicitudes, DynamoDB se usa comúnmente en aplicaciones que requieren mantener el estado de la sesión, en entornos de aplicaciones web sin servidor y en microservicios.

Modelos de datos de Amazon DynamoDB



Consideremos el modelo de datos de Amazon DynamoDB. A diferencia de las bases de datos relacionales, DynamoDB, al ser una base de datos NoSQL, no impone esquemas rígidos.

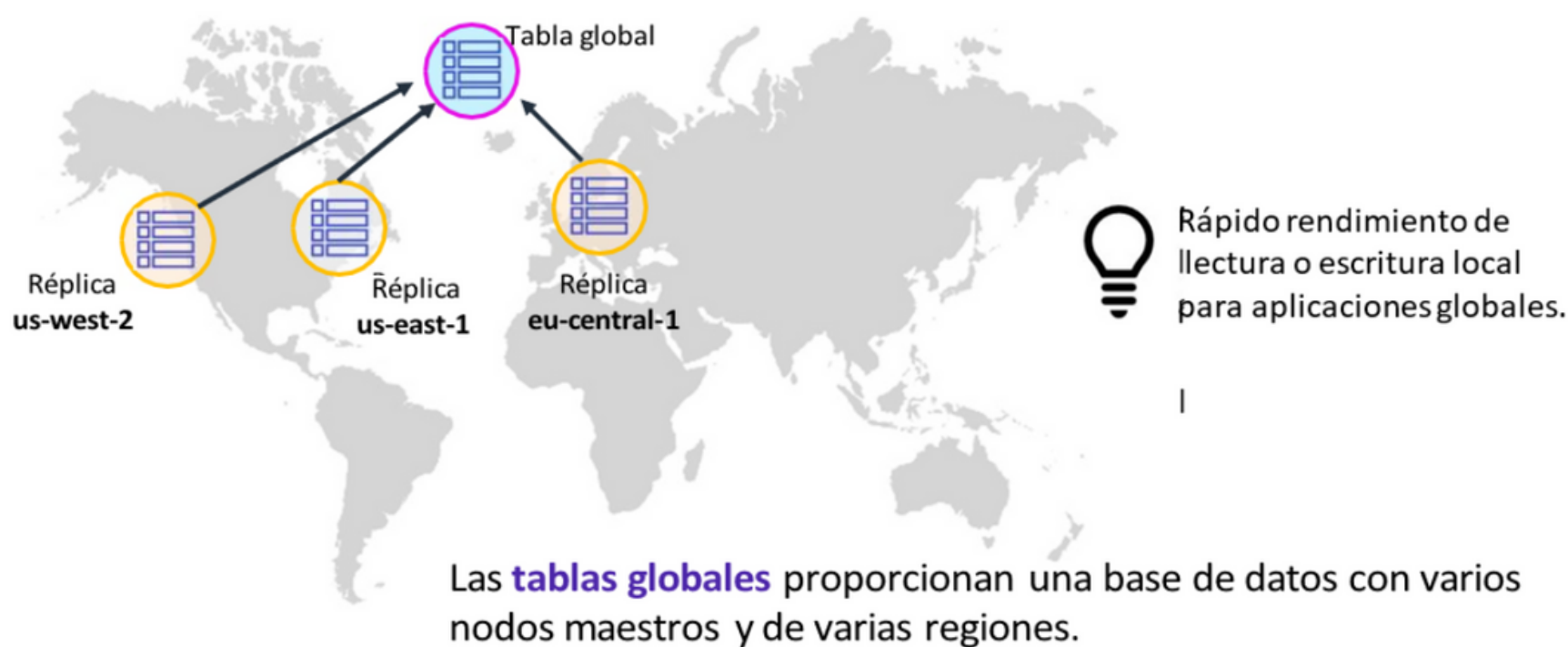
En DynamoDB, una tabla está compuesta por elementos. Cada elemento puede tener un conjunto variable de atributos, donde cada atributo consiste en un par de clave-valor.

Las operaciones básicas en DynamoDB son GET o PUT, que utilizan una clave principal definida por el usuario (conocida como "clave de partición" o "clave hash"). Esta clave principal es esencial para identificar cada elemento en la tabla y se especifica al crear la tabla. Además de estas operaciones, DynamoDB permite consultas flexibles a través de índices secundarios globales e índices secundarios locales, que permiten acceder a atributos no clave primaria.

A medida que el conjunto de datos y la capacidad aprovisionada aumentan, DynamoDB realiza automáticamente particiones para garantizar el rendimiento escalable.

Imaginemos un caso en el que creamos una tabla de DynamoDB para mantener información de pedidos. En esta tabla, cada entrada siempre tendrá un ID de usuario y un ID de pedido. Sin embargo, otros detalles, como la dirección de correo electrónico del usuario, pueden ser recopilados para algunos pedidos, pero no para otros. Conforme se agregan más elementos a la tabla, pueden surgir nuevos atributos. Esta flexibilidad permite a DynamoDB adaptarse a datos que pueden cambiar rápidamente, siendo una excelente opción para datos que son no estructurados, semiestructurados y estructurados.

Tablas globales de Amazon DynamoDB



Por defecto, Amazon DynamoDB replica los datos en múltiples zonas de disponibilidad dentro de una única región. Sin embargo, hay escenarios donde puede necesitar replicar datos en varias regiones.

Las tablas globales de Amazon DynamoDB ofrecen una solución administrada para desplegar una base de datos en múltiples regiones con múltiples nodos maestros, sin necesidad de crear ni mantener su propio sistema de replicación. Al crear una tabla global, especifica las regiones de AWS donde desea que la tabla esté disponible. DynamoDB se encarga de todas las tareas para crear tablas idénticas en estas regiones y propaga los cambios de datos entre ellas.

Por ejemplo, supongamos que tiene una gran base de clientes distribuidos en tres áreas geográficas: costa este de Estados Unidos, costa oeste de Estados Unidos y Europa Occidental. Estos clientes actualizan su información de perfil mientras usan su aplicación.

Para abordar esta necesidad, puede crear una tabla global compuesta por tres tablas CustomerProfiles, cada una específica de una región. DynamoDB automáticamente replicará los cambios de datos entre estas tablas, asegurando que los cambios en los perfiles de cliente en una región se reflejen en las otras. Esto significa que incluso si una región de AWS no está disponible temporalmente, los clientes aún podrán acceder a los mismos datos de perfiles de cliente desde otras regiones.

Caso 1 de Amazon DynamoDB

Marcadores y puntuaciones



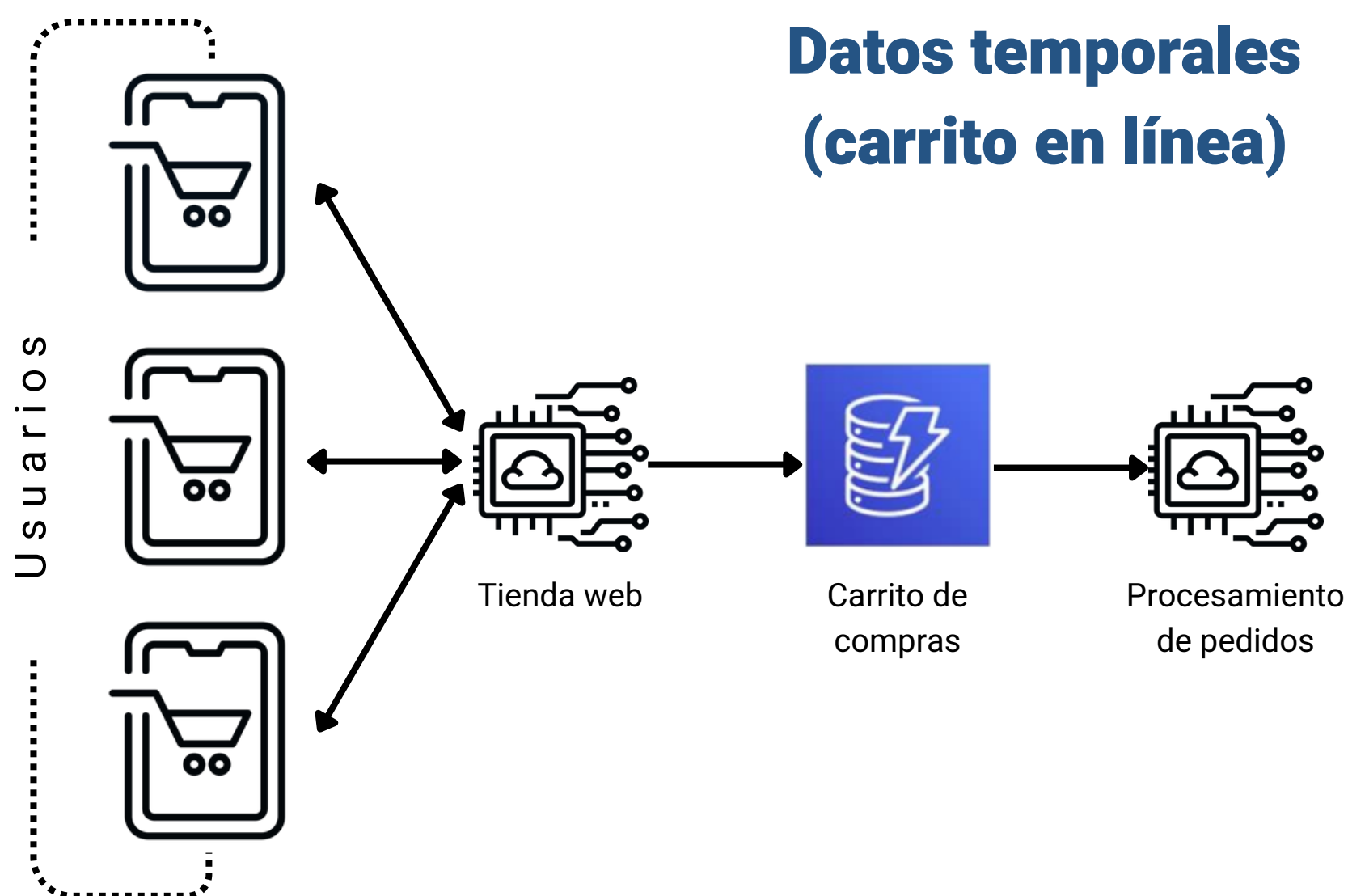
UserId	GameTitle	TopScore	TopScoreDateTime	Victorias	Derrotas	
"101"	"Galaxy Invaders"	5842	"2015-09-15:17:24:31"	21	72	...
"101"	"Meteor Blasters"	1000	"2015-10-22:23:18:01"	12	3	...
"101"	"Starship X"	24	"2015-08-31:13:14:21"	4	9	...
"102"	"Alien Adventure"	192	"2015-07-12:11:07:56"	32	192	...
"102"	"Galaxy Invaders"	0	"2015-09-18:07:33:42"	0	5	...
"103"	"Attack Ships"	3	"2015-10-19:01:13:24"	1	8	...
"103"	"Galaxy Invaders"	2317	"2015-09-11:06:53:00"	40	3	...
"103"	"Meteor Blasters"	723	"2015-10-19:01:13:24"	22	12	...
"103"	"Starship X"	42	"2015-07-11:06:53:00"	4	19	...
...	...	...	...	...	...	...

Las compañías de videojuegos confían en Amazon DynamoDB para gestionar diversos aspectos de sus plataformas, como el estado del juego, los datos de los jugadores, el historial de sesiones y los marcadores. Este servicio les permite escalar de manera confiable para manejar millones de solicitudes y usuarios simultáneos, garantizando al mismo tiempo una baja latencia de forma consistente.

Tomemos como ejemplo el uso de DynamoDB para mantener un marcador de juegos. Los jugadores que están activos en los servidores del juego pueden ver parte de esta información presentada a través de la interfaz de usuario del videojuego. El marcador de GameScores se actualiza continuamente cuando se establecen nuevas puntuaciones máximas o cambian los récords de victorias y derrotas de los principales jugadores. Por ejemplo, Electronic Arts (EA), una destacada empresa de videojuegos con más de 300 millones de jugadores registrados en todo el mundo puede experimentar una alta simultaneidad que se traduce en más de 100,000 solicitudes por segundo y millones de usuarios activos diarios. Consulte Casos de uso de videojuegos de DynamoDB

Caso 2 de Amazon

DynamoDB

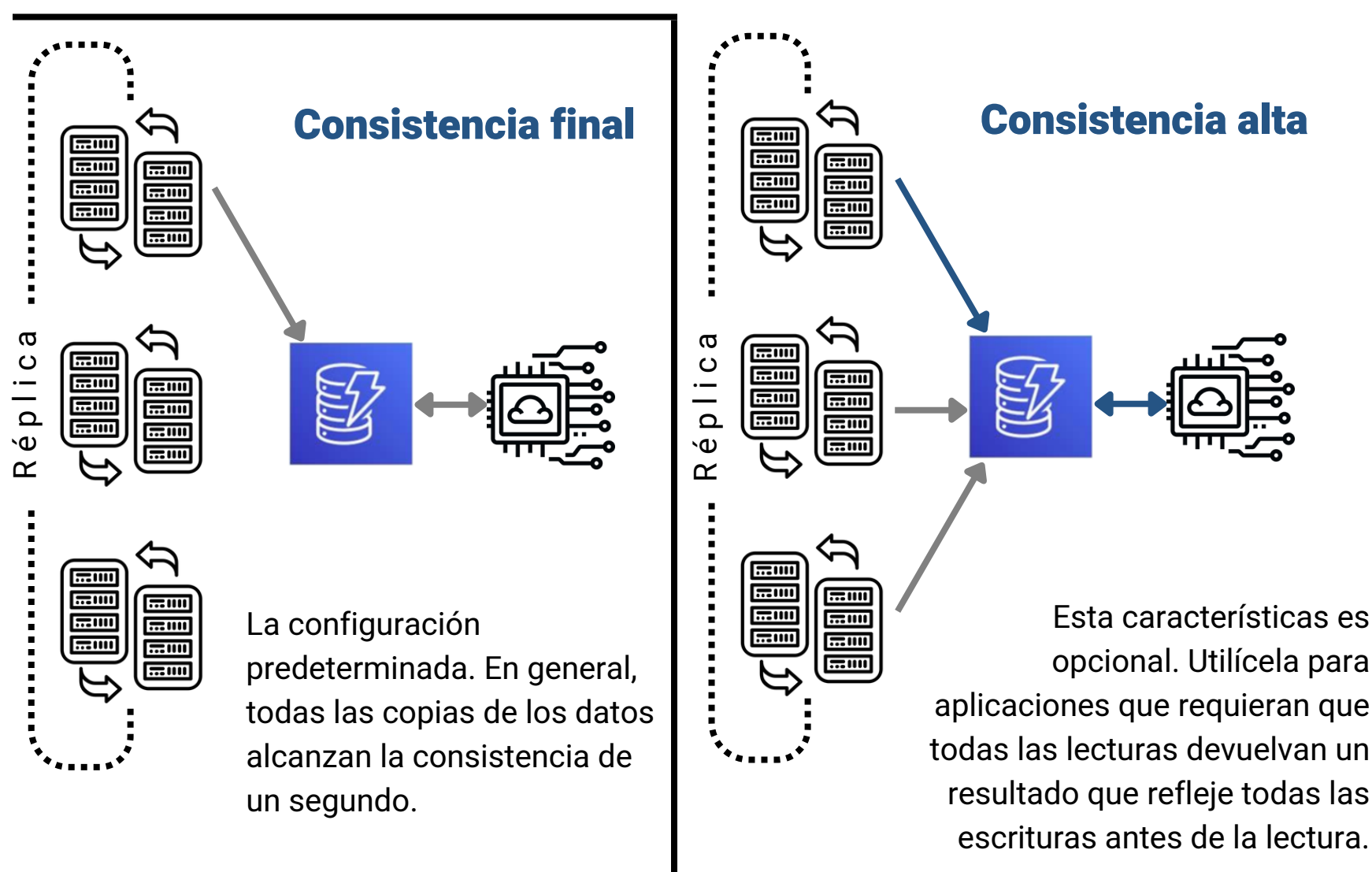


Amazon DynamoDB surgió como respuesta a la necesidad de una base de datos escalable y confiable para impulsar las operaciones esenciales de comercio electrónico de Amazon, como el carrito de compras.

En 2012, Werner Vogels, el CTO de Amazon, explicó los requisitos empresariales del carrito de compras en un artículo de blog al anunciar el lanzamiento de Amazon DynamoDB. Destacó que, para servicios críticos como el carrito de compras, cualquier tiempo de inactividad o degradación del rendimiento tiene un impacto financiero inmediato. Además, se requiere una alta tolerancia a errores y un rendimiento estricto. Asimismo, es esencial la capacidad de escalar gradualmente la infraestructura para adaptarse al crecimiento de las tasas de solicitudes o los tamaños de los conjuntos de datos. También enfatizó la importancia de la predictibilidad en el servicio.

Aunque DynamoDB ha evolucionado con el tiempo, los objetivos fundamentales que lo impulsaron inicialmente siguen siendo relevantes hoy en día para las organizaciones que buscan implementar soluciones de procesamiento para los pedidos del carrito de compras en sus tiendas web.

Opciones de consistencia en Amazon DynamoDB



La consistencia de lectura en Amazon DynamoDB se refiere a cómo y cuándo se reflejan las actualizaciones o inserciones de datos en una operación de lectura posterior para el mismo elemento. DynamoDB le brinda el control para especificar el nivel de consistencia que desea para cada solicitud de lectura en su aplicación.

Las lecturas consistentes finales garantizan que, al leer los datos de una tabla DynamoDB, la respuesta pueda contener información ligeramente desactualizada debido a operaciones de escritura recientes. Sin embargo, si realiza la misma solicitud de lectura después de un breve periodo, la respuesta debería incluir los datos más actualizados.

Por otro lado, las lecturas de alta consistencia devuelven una respuesta que refleja las actualizaciones más recientes de todas las operaciones de escritura previas que se han realizado correctamente. Sin embargo, este tipo de lectura puede no estar disponible si hay retrasos o interrupciones en la red.

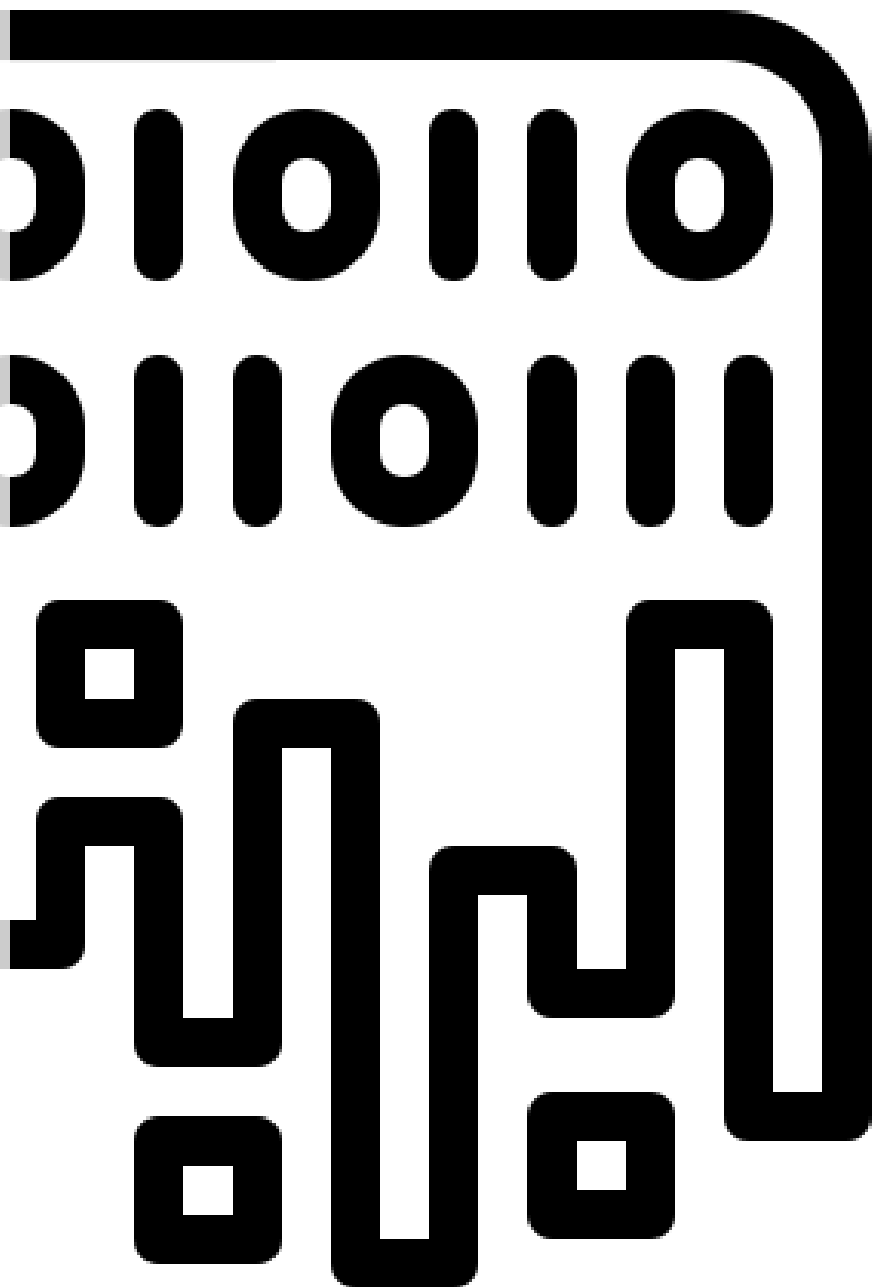
Por defecto, DynamoDB utiliza lecturas consistentes finales, a menos que especifique lo contrario. Las operaciones de lectura como "GetItem", "Query" y "Scan" tienen un parámetro llamado "ConsistentRead". Si establece este parámetro como "true", DynamoDB utilizará lecturas de alta consistencia durante esa operación.



Ahora que conoce las características esenciales clave de Amazon RDS y DynamoDB, considere cuál de estos dos servicios de base de datos sería la mejor opción para el caso de uso de la cafetería.

En una unidad anterior del Bootcamp, completó un laboratorio de desafíos donde implementó una base de datos MySQL que se ejecuta en una instancia EC2. Esta base de datos aloja una tabla que contiene detalles sobre el menú de la cafetería. También proporciona la capa de almacenamiento de datos para los pedidos que los clientes realizan en línea.

Si el equipo de la cafetería decide que quiere migrar la capa de base de datos para poder utilizar un servicio de base de datos administrada de Amazon, ¿cuál de estas dos opciones de base de datos sería la mejor opción?



El mentor dirigirá la conversación en la clase sobre este tema. Lo animamos a participar en el diálogo y explicar el razonamiento detrás de su respuesta.

Estos son algunos de los aprendizajes clave de esta lección:

- Amazon DynamoDB es un servicio de base de datos NoSQL no relacional completamente administrada de clave-valor y documentos.
- DynamoDB no tiene servidor, proporciona escalado horizontal extremo y baja latencia.
- Las tablas globales de DynamoDB garantizan que los datos se repliquen en varias regiones.
- DynamoDB proporciona consistencia final de forma predeterminada (en general, es totalmente consistente para lecturas un segundo después de la escritura). La consistencia alta también es una opción.

