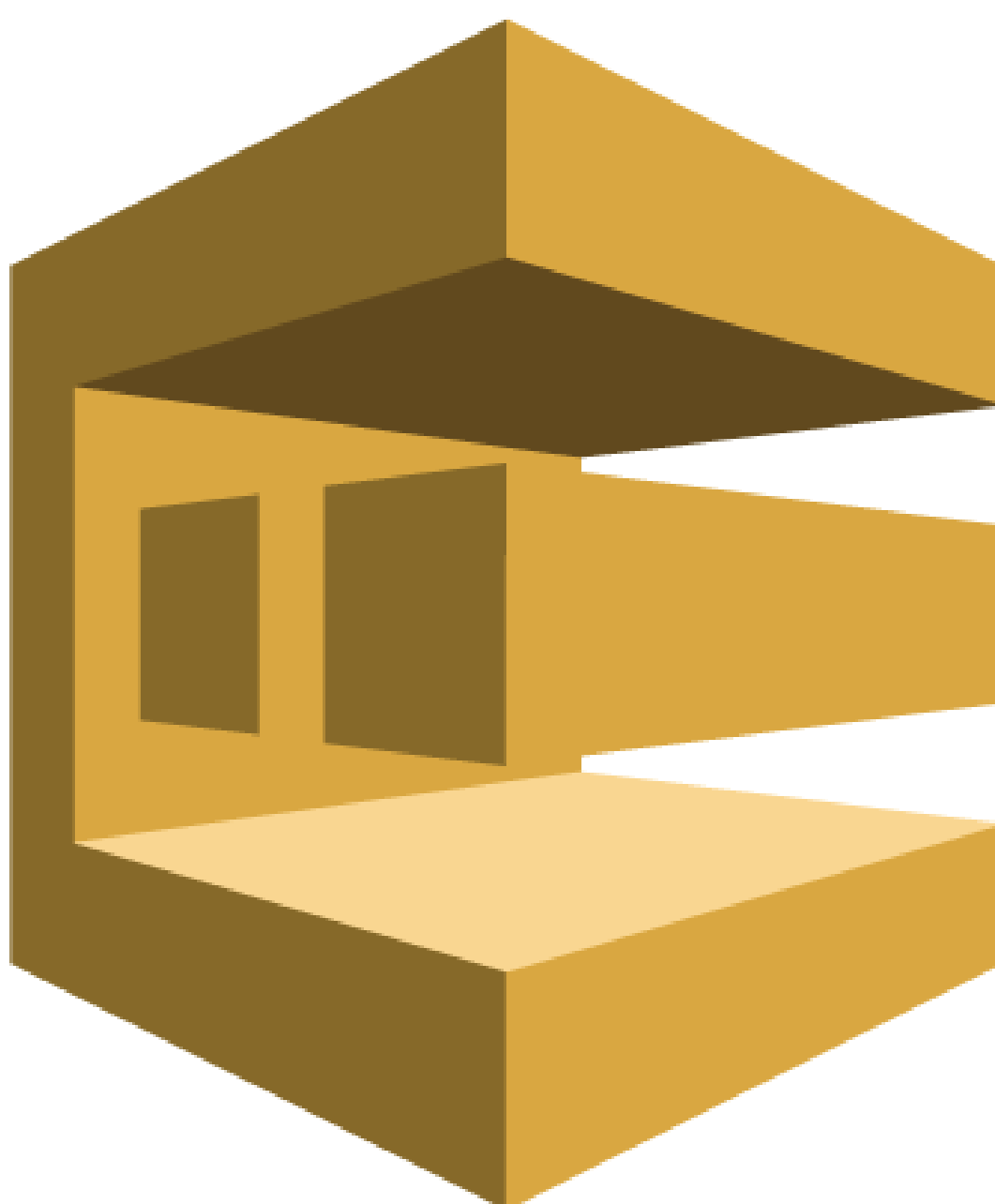


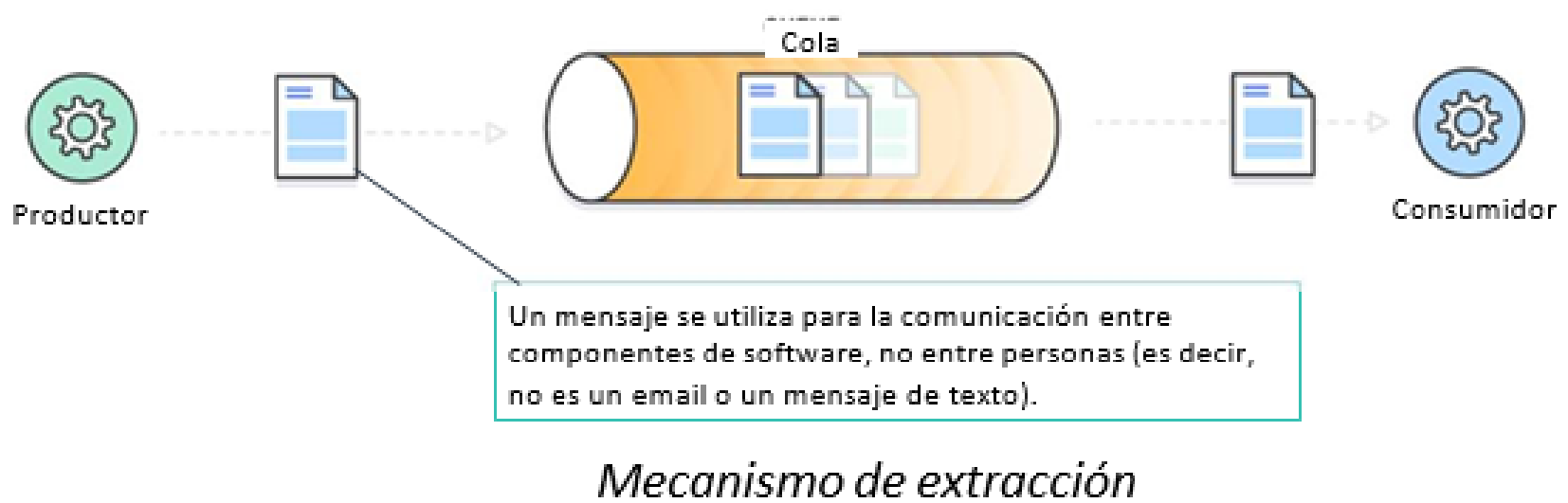
LECCIÓN 3: DESACOPLAMIENTO CON AMAZON SQS.



COLAS DE MENSAJES PARA DESACOPLAR ARQUITECTURAS

Productor: componente de la aplicación que produce mensajes y los agrega a la cola

Consumidor: componente de la aplicación que sondea la cola de mensajes y los procesa



Las colas de mensajes son otro componente que lo ayuda a implementar una arquitectura desacoplada. Las colas de mensajes proporcionan comunicación y coordinación para estas aplicaciones distribuidas.

Una cola de mensajes es un repositorio temporal de mensajes que están esperando ser procesados. Los mensajes suelen ser pequeños y pueden ser cosas como solicitudes, respuestas, mensajes de error o información. Ejemplos de mensajes incluyen registros de clientes, pedidos de productos, facturas, registros de pacientes, etc.

Para enviar un mensaje, un componente llamado productor agrega un mensaje a la cola. El mensaje se almacena en la cola hasta que otro componente llamado consumidor lo recupera y procesa.

AMAZON SQS

- **Servicio de cola de mensajes completamente administrado**
 - **Utiliza un mecanismo de extracción**
- **Los mensajes se cifran y almacenan hasta que se procesan y eliminan**
- **Funciona como un búfer entre productores y consumidores**



Amazon Simple
Queue Service
(Amazon SQS)

Amazon Simple Queue Service (Amazon SQS) es un servicio de cola de mensajes completamente administrado que le permite desacoplar los componentes de la aplicación para que se ejecuten de forma independiente. Permite a las aplicaciones de servicios web poner en cola los mensajes que genera un componente de la aplicación para que los consuma otro componente.

Una cola es un repositorio temporal de mensajes que están esperando ser procesados. Los mensajes se almacenan hasta que se procesan y eliminan (de 1 a 14 días; la cantidad predeterminada es 4 días). Los mensajes pueden contener hasta 256 KB de texto en cualquier formato. Amazon SQS funciona a muy gran escala y procesa miles de millones de mensajes al día. Almacena todas las colas de mensajes y los mensajes dentro de una región única de AWS de alta disponibilidad con varias zonas de disponibilidad redundantes. Ningún error en el equipo, la red o la zona de disponibilidad puede hacer que los mensajes sean inaccesibles. Los mensajes pueden enviarse y leerse de manera simultánea.



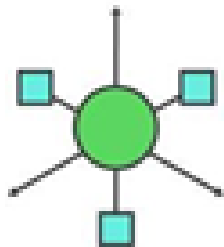
Usted puede compartir de manera segura colas de Amazon SQS anónimamente o con cuentas de AWS específicas. También puede restringir el uso compartido de colas por dirección IP y por hora del día. Los mensajes de las colas de SQS están cifrados con cifrado del lado del servidor (SSE) mediante claves administradas en AWS Key Management Service (AWS KMS). Amazon SQS descifra los mensajes únicamente cuando se envían a un consumidor autorizado.

Amazon SQS admite que varios productores y consumidores interactúen con la misma cola. Amazon SQS se puede utilizar con varios servicios de AWS, entre ellos:

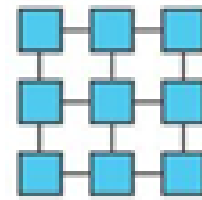
Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3), Amazon Elastic Container Service (Amazon ECS), AWS Lambda y Amazon DynamoDB.

LOGRAR BAJO ACOPLAMIENTO CON AMAZON SQS

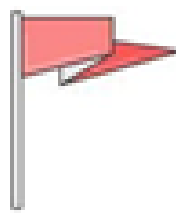
CON AMAZON SQS, PUEDE HACER LO SIGUIENTE:



Utilizar el **procesamiento asincrónico** para obtener las respuestas de cada paso rápidamente



Manejar los **requisitos de rendimiento y servicio** aumentando la cantidad de instancias de trabajo



Recuperarse **fácilmente de los pasos erróneos**, ya que los mensajes permanecen en la cola

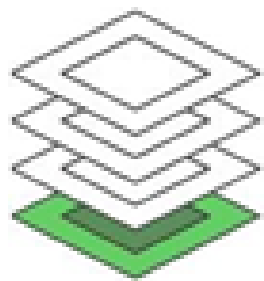
Amazon SQS le permite lograr bajo acoplamiento en su arquitectura.

Realiza un procesamiento asincrónico para que pueda obtener respuestas de cada paso rápidamente.

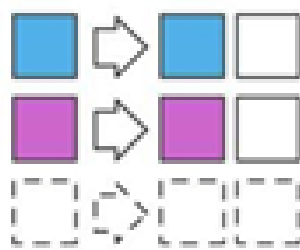
Puede manejar los requisitos de rendimiento y servicio aumentando la cantidad de instancias de trabajo.

Su aplicación puede recuperarse fácilmente de pasos erróneos, ya que los mensajes permanecen en la cola.

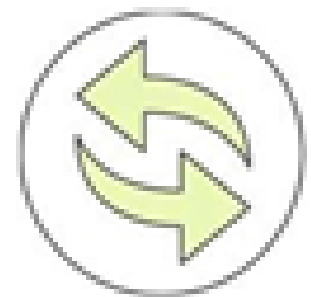
CASOS DE USO GENERALES DE AMAZON



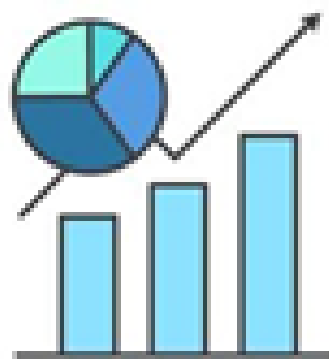
Colas de trabajo



Operaciones por lotes de búfer



Descarga de solicitudes



Activar Amazon EC2 Auto Scaling

Amazon SQS se puede utilizar de muchas maneras diferentes:

COLAS DE TRABAJO:

Desacople los componentes de una aplicación distribuida que puede que no procesen todos la misma carga de trabajo al mismo tiempo.

OPERACIONES POR LOTES DE BÚFER:

Agregue escalabilidad y fiabilidad a la arquitectura y atenúe los picos de volumen temporales sin perder mensajes y sin aumentos de latencia.

DESCARGA DE SOLICITUDES:

Ponga solicitudes en cola para retirar las operaciones lentas de las rutas de solicitud interactivas.

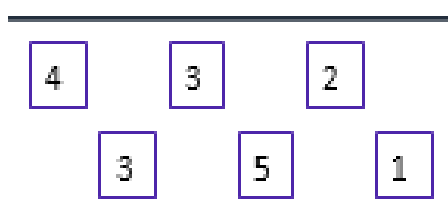
DESENCADENAMIENTO DE AMAZON EC2 AUTO SCALING:

Utilice colas de SQS para ayudar a determinar la carga en una aplicación.

Cuando se combinan con Amazon EC2 Auto Scaling, puede escalar el número de instancias EC2 en forma ascendente o descendente, según el volumen de tráfico.

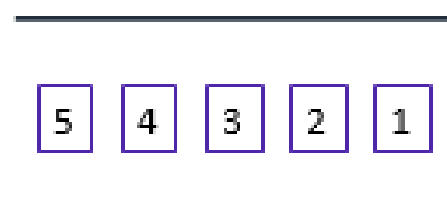
TIPOS DE COLA

Colas estándar



- Se realizan entregas **al menos una vez**
- Orden **para optimizar el esfuerzo**
- Rendimiento **casi ilimitado**

Colas en que el primero en entrar es el primero en salir (FIFO)



- Entrega en que **el primero en entrar es el primero en salir**
- Procesamiento **por única vez**
- **Alto** rendimiento

EXISTEN DOS TIPOS DE COLAS DE SQS:

Las colas estándar ofrecen lo siguiente:

Al menos una entrega:
los mensajes se entregan al menos una vez, aunque en ocasiones se entregan varias copias del mensaje.

Orden para optimizar el esfuerzo:
en ocasiones, los mensajes se entregan en un orden distinto al que se enviaron.

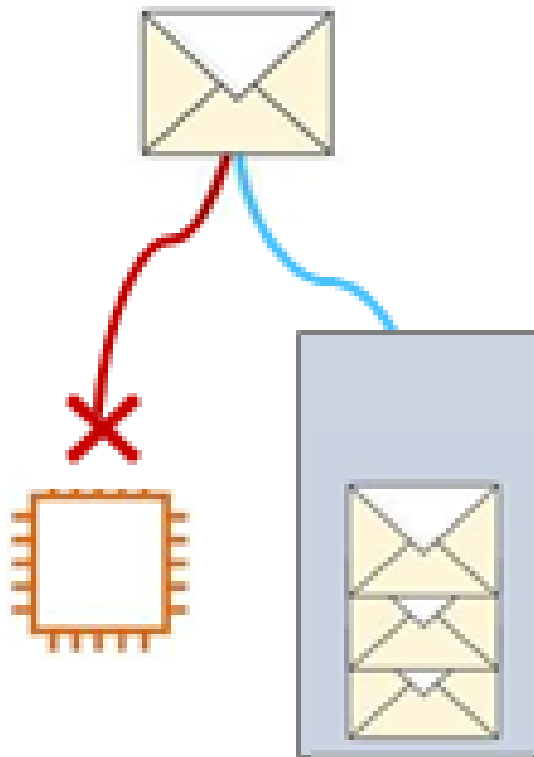
Capacidad de procesamiento casi ilimitada:
las colas estándar admiten un número casi ilimitado de transacciones por segundo (TPS) para cada acción de la API.

Colas en que el primero en entrar es el primero en salir (FIFO):

Están diseñadas para garantizar que los mensajes se procesen exactamente una vez, en el orden exacto en el que se enviaron.

Capacidad de procesamiento alta:
Las colas FIFO admiten hasta 300 mensajes por segundo (300 operaciones de envío, recepción o eliminación por segundo). Si agrupa 10 mensajes por operación (máximo), las colas FIFO pueden admitir hasta 3000 mensajes por segundo.

CARACTERÍSTICAS DE AMAZON SQS (1 DE 3)



Compatibilidad con cola de mensajes fallidos:

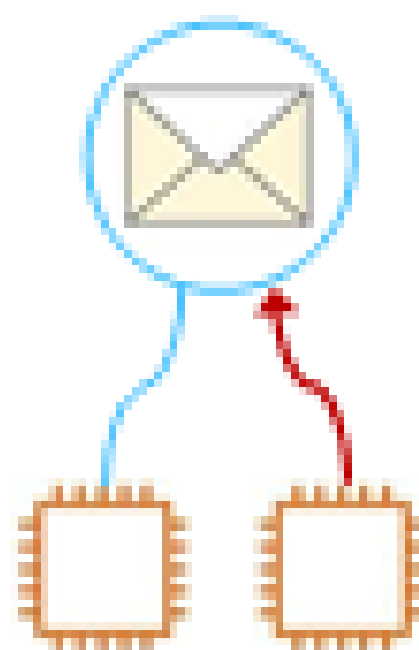
Amazon SQS ofrece compatibilidad con colas de mensajes fallidos. Una cola de mensajes fallidos (DLQ) es una cola de mensajes que no pudieron procesarse. Recibe los mensajes una vez que se alcanza el número máximo de intentos de procesamiento. Una DLQ es como cualquier otra cola de SQS: los mensajes se envían y se reciben. Puede crear una DLQ desde la API de Amazon SQS y la consola.



CARACTERÍSTICAS DE AMAZON SQS (2 DE 3)

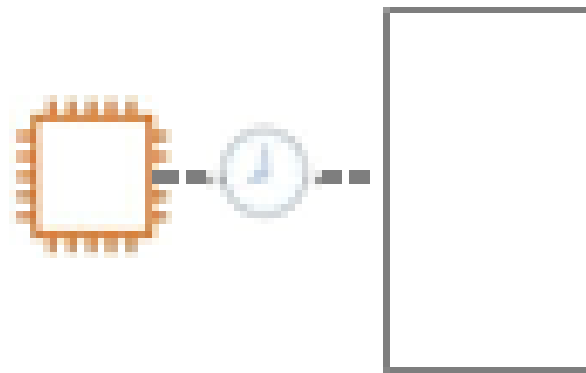
Tiempo de espera de visibilidad:

Otra característica de Amazon SQS es el tiempo de espera de visibilidad. El tiempo de espera de visibilidad es el intervalo en el que Amazon SQS impide que otros consumidores reciban y procesen el mismo mensaje. El tiempo de espera ayuda a garantizar que un trabajo no se procese varias veces y ocasione una duplicación. Durante el tiempo de espera de visibilidad, el componente que recibió el mensaje lo procesa y, a continuación, lo elimina de la cola. El tiempo de espera de visibilidad predeterminado para un mensaje es de 30 segundos y el máximo es de 12 horas.



Si el consumidor no procesa y elimina el mensaje antes de que expire el tiempo de espera de visibilidad, el mensaje se vuelve visible para otros consumidores y podría procesarse de nuevo. Normalmente, debe establecer el tiempo de espera de visibilidad como el tiempo máximo que tarda su aplicación en procesar y eliminar un mensaje de la cola.

SONDEO LARGO



Por último, Amazon SQS admite sondeo corto y sondeo largo para recuperar mensajes de las colas de SQS. De forma predeterminada, las colas usan sondeo corto.

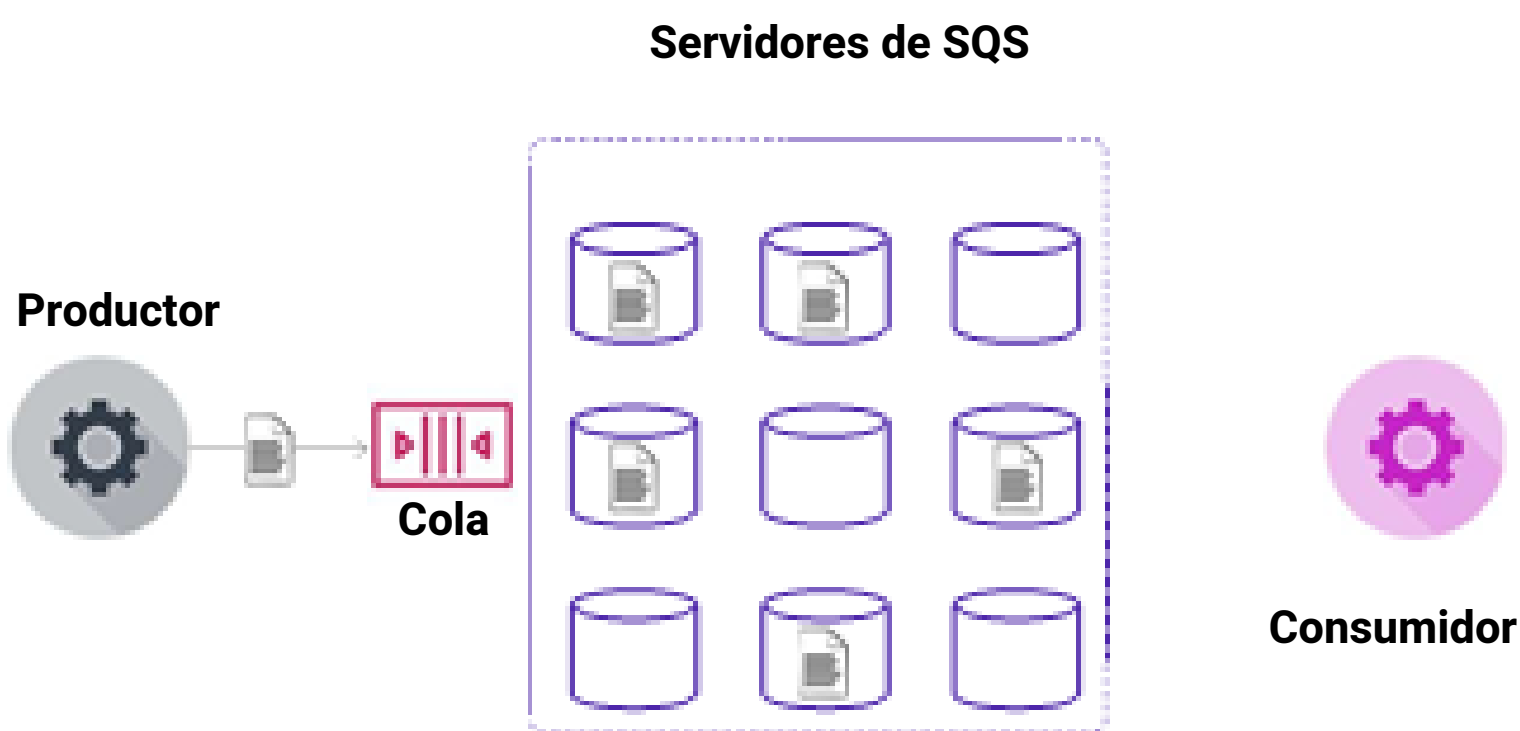
El sondeo corto consulta solo un subconjunto de los servidores (en función de una distribución aleatoria ponderada) para encontrar mensajes que puedan incluirse en la respuesta. Amazon SQS envía la respuesta inmediatamente, incluso si la consulta no encontró ningún mensaje.

Por el contrario, el sondeo largo consulta todos los servidores en busca de mensajes. Amazon SQS envía la respuesta después de que recopila el número máximo de mensajes para la respuesta o de que expira el tiempo de espera del sondeo.

El sondeo largo permite la recuperación de los mensajes de la cola de mensajes de SQS tan pronto como están disponibles. El sondeo largo puede reducir el costo de uso de Amazon SQS, ya que se reduce el número de recepciones vacías.

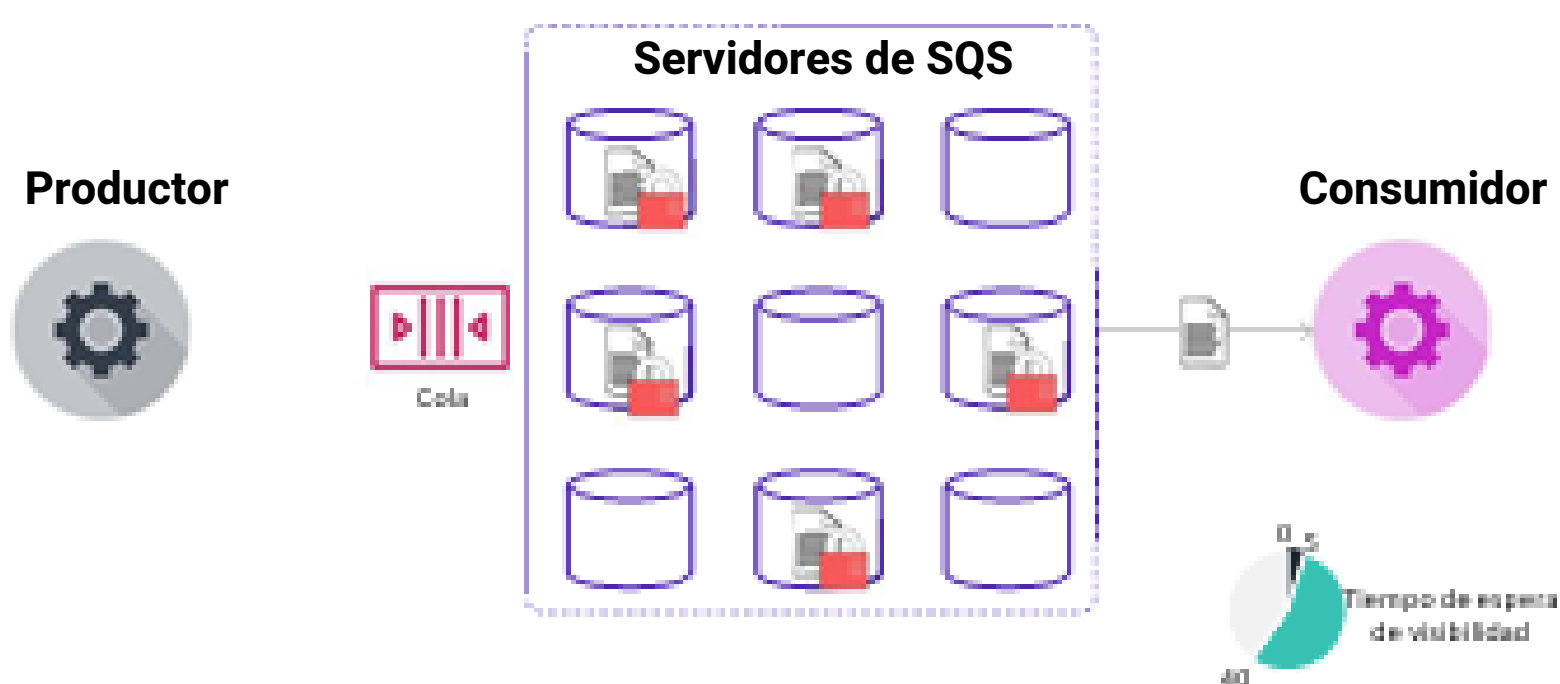
CICLO DE VIDA DE LOS MENSAJES DE AMAZON SQS: CREACIÓN

El ciclo de vida de un mensaje en una cola de SQS se puede ilustrar con el siguiente caso. Primero, un productor envía un mensaje a la cola y el mensaje se distribuye entre los servidores de SQS de forma redundante.



**El productor envía un mensaje
a la cola, que se distribuye de
forma redundante**

CICLO DE VIDA DE LOS MENSAJES DE AMAZON SQS: PROCESO

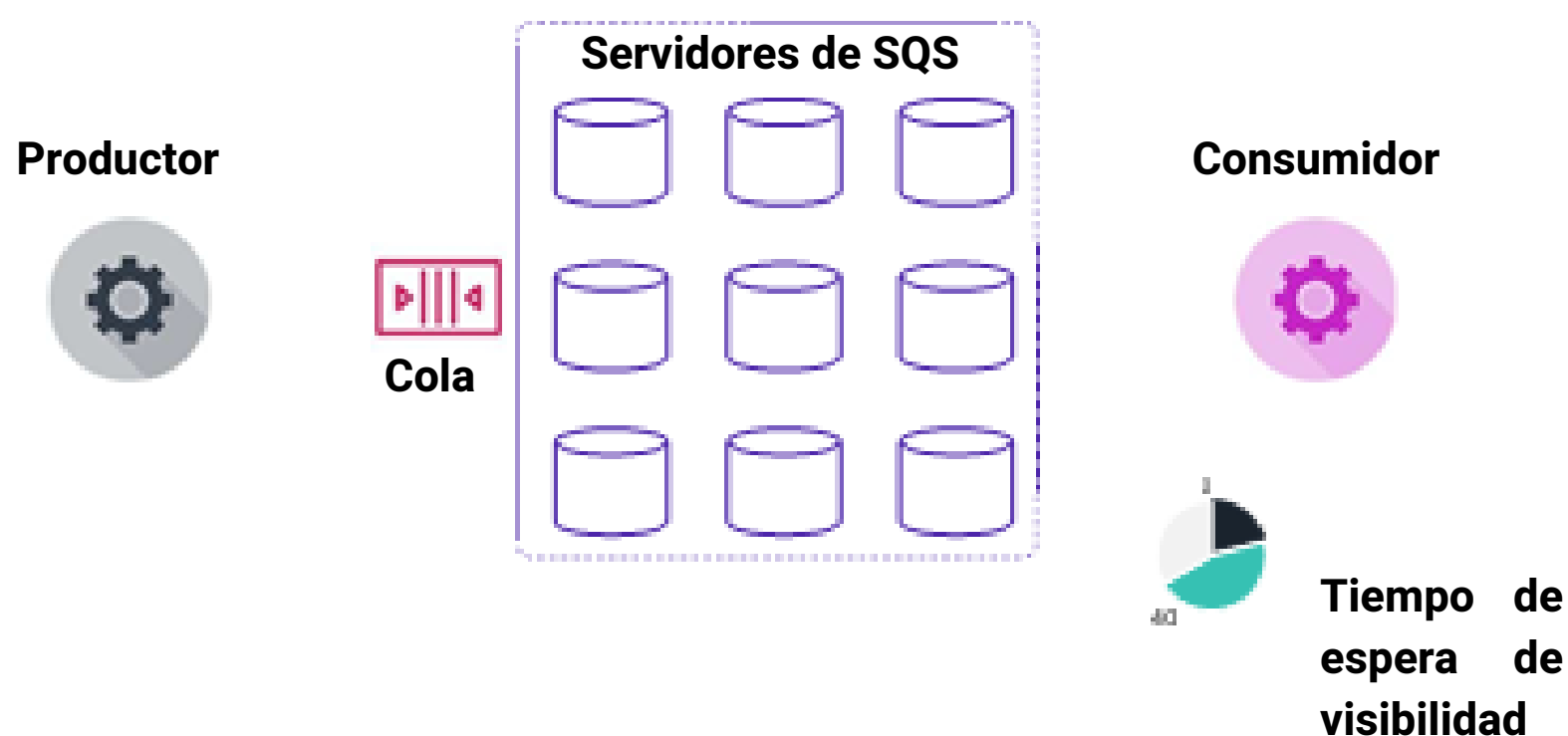


**El mensaje se recoge para que
un consumidor lo procese y
comienza el tiempo de espera
de visibilidad**

Cuando un consumidor está listo para procesar el mensaje, recupera el mensaje de la cola. Mientras el mensaje se procesa, permanece en la cola.

Durante el tiempo de espera de visibilidad, otros consumidores no pueden procesar el mensaje. En este ejemplo, el tiempo de espera de visibilidad es de 40 segundos.

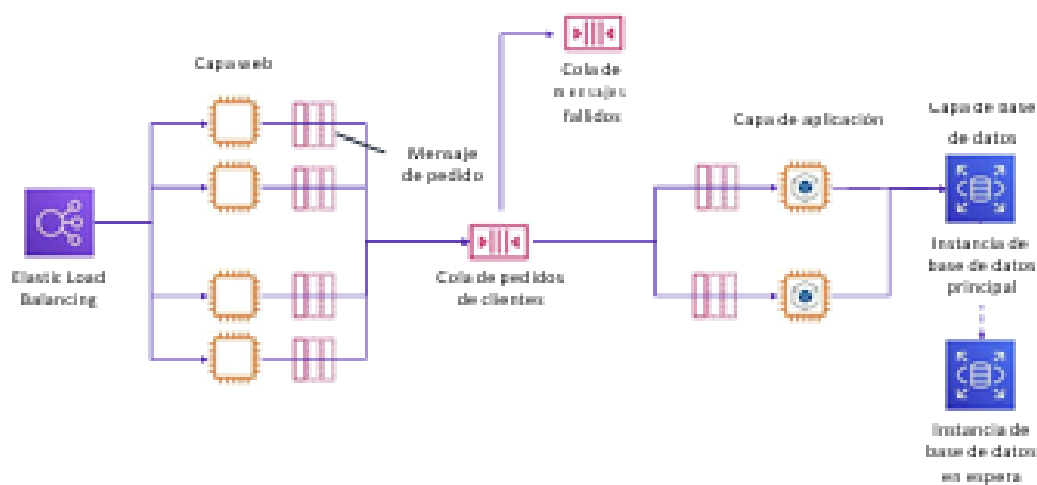
CICLO DE VIDA DE LOS MENSAJES DE AMAZON SQS: ELIMINACIÓN



Después de procesar el mensaje, el consumidor elimina el mensaje de la cola. Esta acción evita que el mensaje se reciba y procese de nuevo cuando expire el tiempo de espera de visibilidad.

Amazon SQS no elimina automáticamente el mensaje. Dado que Amazon SQS es un sistema distribuido, no hay garantías de que el consumidor reciba realmente el mensaje (por ejemplo, debido a un problema de conectividad o a un problema en la aplicación del consumidor). Por lo tanto, el consumidor debe eliminar el mensaje de la cola después de recibirlo y procesarlo.

EJEMPLO DE DESACOPLAMIENTO: USO DE AMAZON SQS



Nuevamente, considere la aplicación de procesamiento de pedidos sobre la que aprendió en la Lección anterior. Puede desacoplar la arquitectura de esta aplicación introduciendo una cola de SQS. Puede utilizar la cola para aislar la lógica de procesamiento en su propio componente para que se ejecute en un proceso separado de la aplicación web.

Este diseño permite que el sistema sea más resistente a los picos de tráfico. Además, permite que el trabajo se realice tan rápido como sea necesario para administrar los costos.

Además, ahora tiene un mecanismo para almacenar los pedidos como mensajes (con la cola que funciona como una base de datos temporal). También movió el alcance de su transacción con su base de datos más abajo en la pila. Si se produce un error en la excepción de aplicación o transacción, este diseño ayuda a garantizar que el procesamiento de pedidos pueda retirarse o redirigirse a una cola de mensajes fallidos para su reprocesamiento en una etapa posterior.

CASOS DE USO DE LA COLA DE MENSAJES

Estos casos de uso comunes demuestran dónde es muy adecuada una cola de mensajes:



Comunicación de
servicio a servicio



Elementos de trabajo
asíncronos



Notificaciones de
cambio de estado

Comunicación de servicio a servicio:

Por ejemplo, considere un sitio web del frontend que debe actualizar la dirección de entrega de un cliente en un servicio de gestión de relaciones con clientes (CRM) del backend. Puede hacer que el código del sitio web del frontend envíe mensajes a una cola de SQS y que el servicio CRM del backend los consuma.

Elementos de trabajo asíncronos:

por ejemplo, supongamos que el sistema de reserva de un hotel debe cancelar una reserva, que es un proceso que toma mucho tiempo. Puede enviar mensajes a una cola de SQS y hacer que el mismo sistema de reserva del hotel consuma esos mensajes y realice cancelaciones asíncronas.

Notificaciones de cambio de estado:

Por ejemplo, supongamos que tiene un servicio que administra algún recurso. Quiere que otros servicios reciban actualizaciones sobre los cambios realizados en ese recurso. Un sistema de inventario podría publicar notificaciones cuando las existencias de un elemento determinado son bajas y es necesario hacer un pedido.

CASOS DE USO DE LA COLA DE MENSAJES

 Comunicación de servicio a servicio	 Selección de mensajes específicos
 Elementos de trabajo asíncronos	 Mensajes grandes
 Notificaciones de cambio de estado	

También es importante saber cuándo una tecnología en particular no se ajusta bien a su caso de uso. La mensajería tiene su propio conjunto de antipatrones habituales.

Selección de mensajes específicos:

Es posible que desee recibir selectivamente mensajes de una cola que coincidan con un determinado conjunto de atributos o con una consulta lógica ad hoc. Por ejemplo, un servicio solicita un mensaje con un atributo determinado porque contiene una respuesta a otro mensaje enviado por el servicio. En este caso, la cola puede tener mensajes que nadie sondea y que nunca se consumen.

Mensajes grandes:

La mayoría de los protocolos y las implementaciones de mensajería funcionan mejor con mensajes de tamaño razonable (en decenas o cientos de KB). A medida que aumentan los tamaños de los mensajes, es mejor utilizar un sistema de almacenamiento exclusivo (como Amazon S3) y pasar una referencia a un objeto del almacén que es el mensaje propiamente dicho.

ESTOS SON ALGUNOS DE LOS APRENDIZAJES CLAVE DE ESTA LECCIÓN:

- **Amazon SQS es un servicio de cola de mensajes completamente administrado que le permite desacoplar los componentes de la aplicación para que se ejecuten de forma independiente.**
- **Amazon SQS admite las colas estándar y las colas FIFO.**
- **Un productor envía un mensaje a una cola. Un consumidor procesa y elimina el mensaje durante el tiempo de espera de visibilidad.**
- **Los mensajes que no se pueden procesar pueden enviarse a una cola de mensajes fallidos.**
- **El sondeo largo es una forma de recuperar una gran cantidad de mensajes de las colas de SQS.**

