COLOMBIA POTENCIA DE LA VIDA

TIC

Lección 4
Almacenamiento de sesiones web en caché

TALENTO TECH

AZ | PROYECTOS EDUCATIVOS

UTP Universidad Tecnológica de Pereira

Administración de sesiones: sesiones rápidas

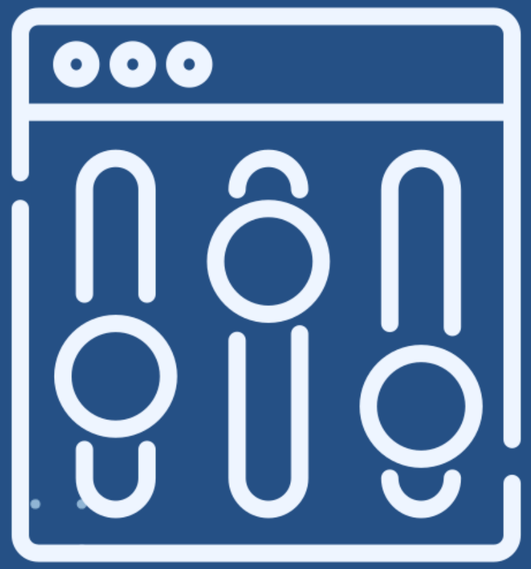
Cuando un usuario o un servicio interactúan con una aplicación web, envían una solicitud HTTP y la aplicación devuelve una respuesta. Una secuencia de tales transacciones se denomina sesión. Cada solicitud es independiente de las transacciones anteriores. Por lo tanto, las sesiones se utilizan para administrar la autenticación de los usuarios y almacenar sus datos mientras interactúan con la aplicación. Por ejemplo, al emplear sesiones, los usuarios no tienen que enviar sus credenciales por cada solicitud que realicen al servidor.

Elastic Load Balancing

Sesiones rápidas

Es una característica que permite a un balanceador de carga dirigir una solicitud al servidor que administra la sesión del usuario específicamente.

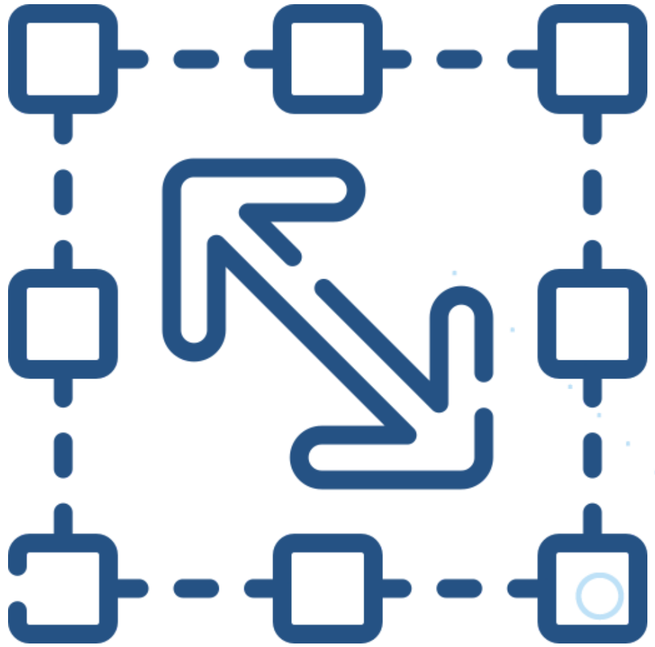
- Utilizan cookies del lado del cliente
- Son rentables
- Aceleran la recuperación de sesiones
- Tienen desventajas:
  - Pérdida de sesiones cuando se produce un error en la instancia
  - Escalabilidad limitada: distribución desigual de la carga y mayor latencia



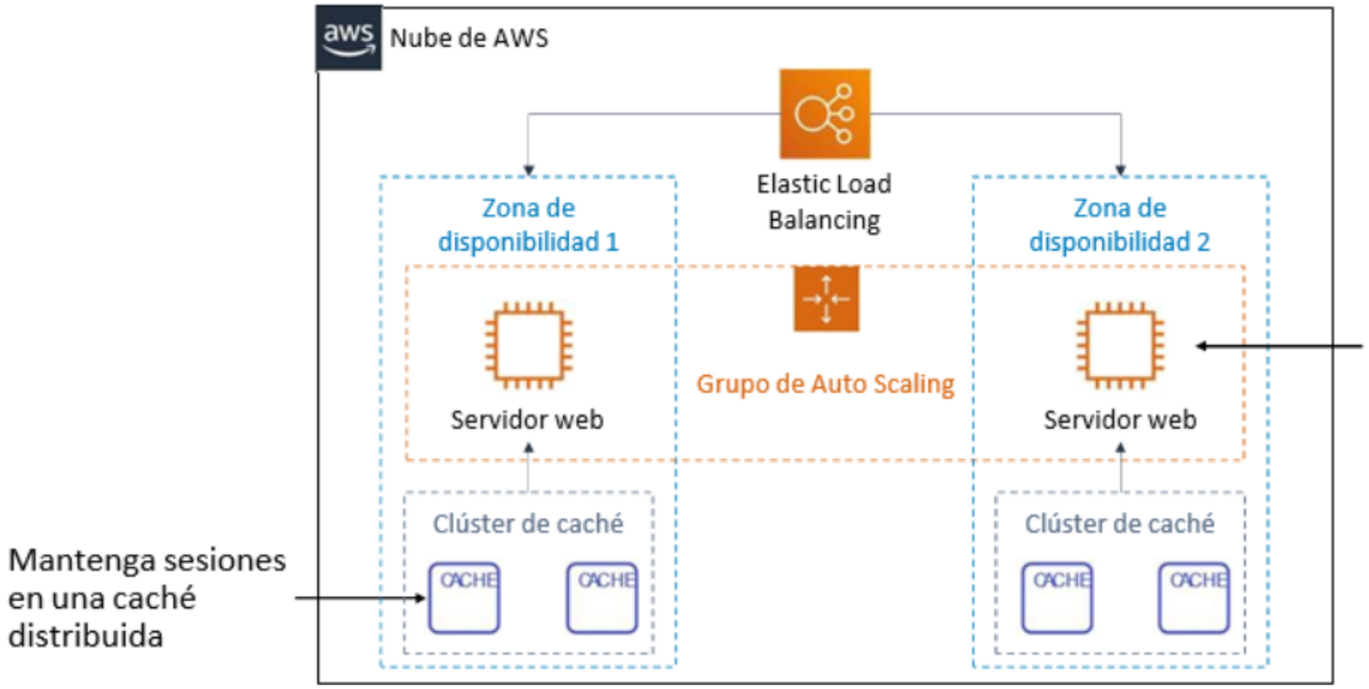
Puede administrar las sesiones de los usuarios de distintas maneras. (De forma predeterminada, un balanceador de carga dirige cada solicitud independientemente de la instancia registrada con la carga más pequeña). Para utilizar las sesiones rápidas, los clientes deben admitir las cookies.

Las sesiones rápidas son rentables porque se almacenan en los servidores web que ejecutan sus aplicaciones. Por lo tanto, eliminan la latencia de la red y aceleran la recuperación de esas sesiones. Sin embargo, en caso de error en la instancia, es probable que pierda las sesiones almacenadas en esa instancia.

Otra desventaja de las sesiones rápidas es que pueden limitar la escalabilidad de su aplicación. Con las sesiones rápidas, el balanceador de carga no puede equilibrar realmente la carga cada vez que recibe una solicitud de un cliente. Con las sesiones rápidas, el balanceador de carga se ve obligado a enviar todas las solicitudes al servidor original en el que se creó el estado de la sesión. Si ese servidor está muy cargado, la recepción de tantas solicitudes puede provocar una carga desigual entre los servidores y afectar el tiempo de respuesta del usuario.

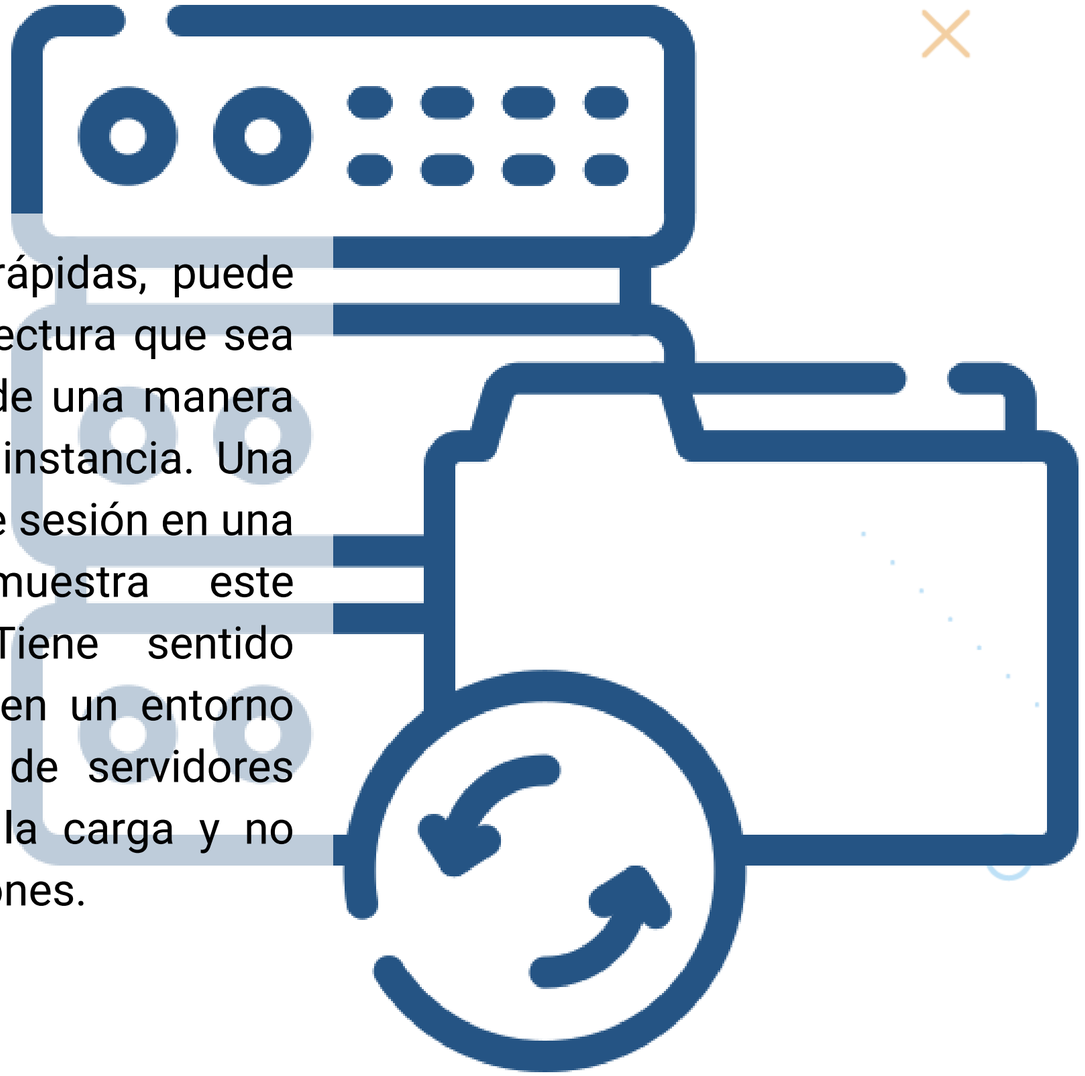


En lugar de sesiones rápidas: mantenga las sesiones dentro de una caché distribuida



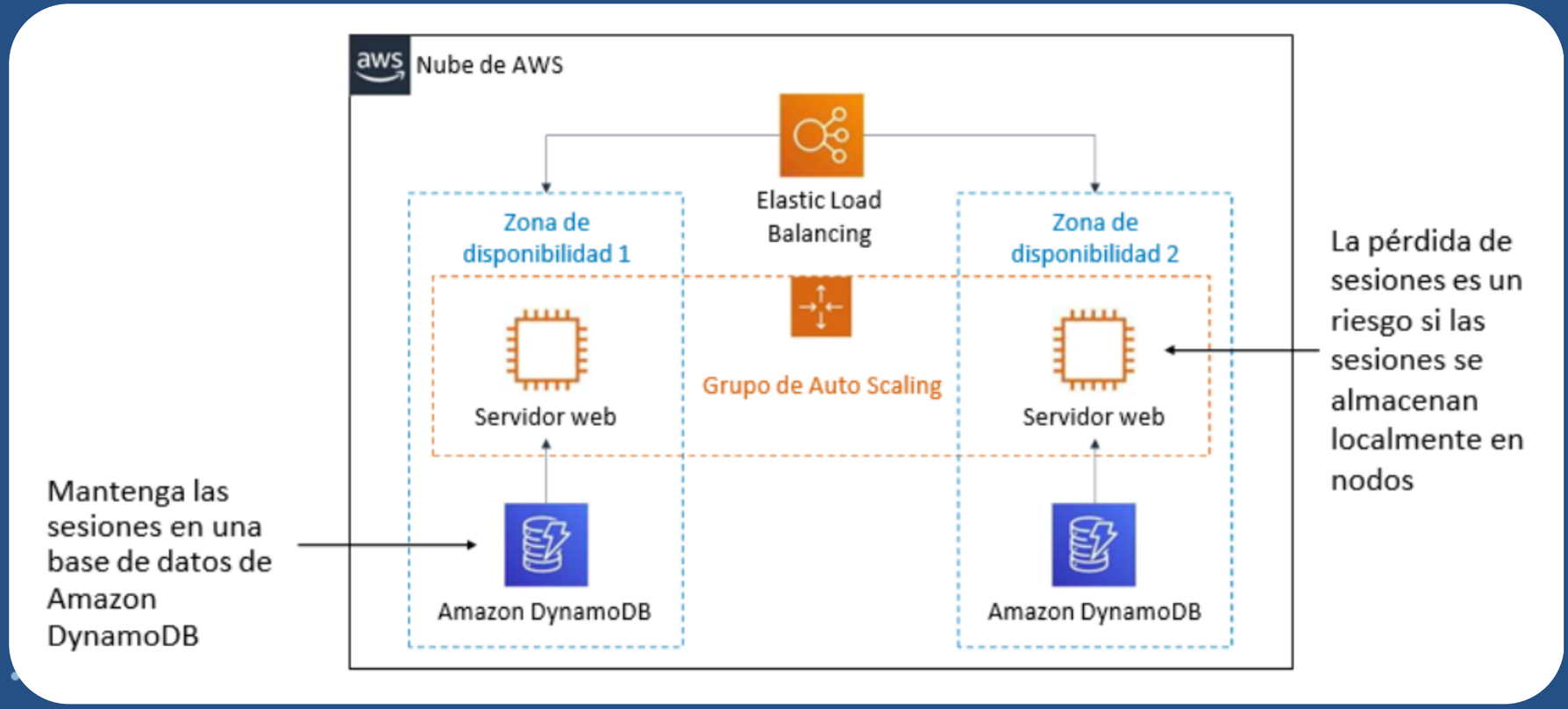
Mantenga sesiones en una caché distribuida

La pérdida de sesiones es un riesgo si las sesiones se almacenan localmente en nodos



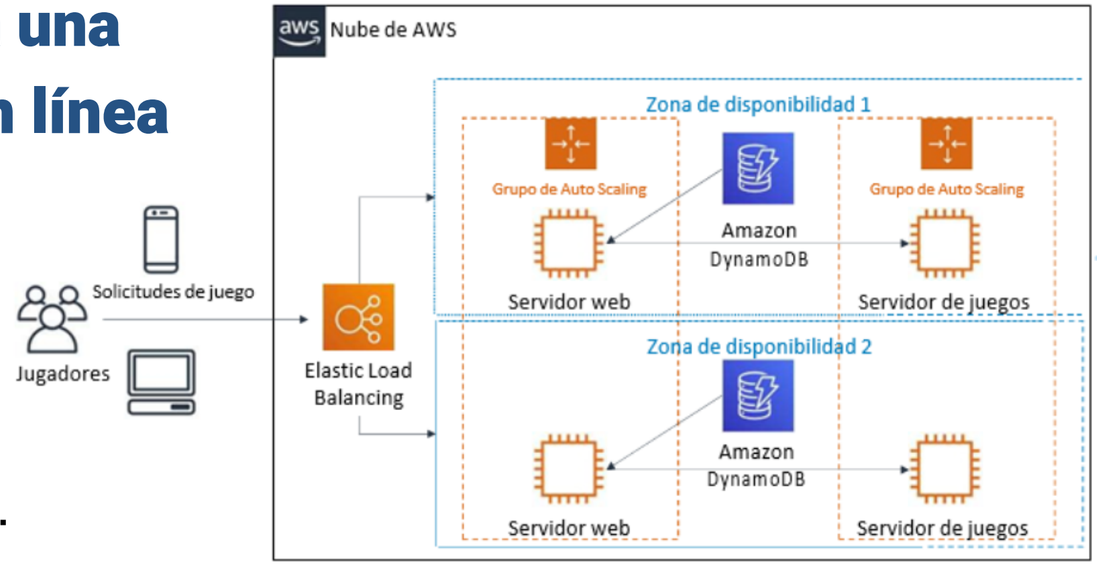
En lugar de utilizar sesiones rápidas, puede designar una capa en su arquitectura que sea capaz de almacenar sesiones de una manera escalable y sólida fuera de la instancia. Una opción es mantener los datos de sesión en una caché distribuida, como muestra este diagrama de arquitectura. Tiene sentido implementar esta arquitectura en un entorno dinámico cuando la cantidad de servidores web cambia para ajustarse a la carga y no desea arriesgarse a perder sesiones.

En lugar de sesiones rápidas: mantenga las sesiones dentro de una tabla de DynamoDB

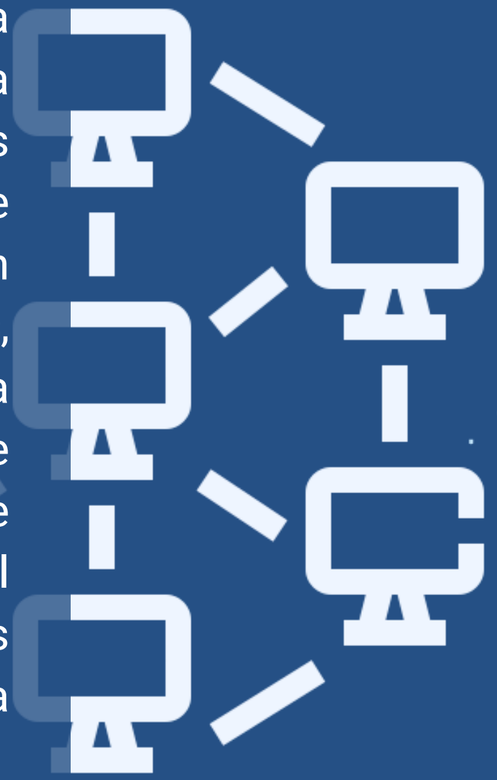


Otra opción consiste en mantener los datos de sesión en una base de datos de Amazon DynamoDB, como se muestra en este diagrama. Con Amazon DynamoDB, puede establecer su política de escalado y hacer que DynamoDB aumente o reduzca su capacidad según sea necesario.

Ejemplo: almacenamiento de estados de sesión para una aplicación de juegos en línea



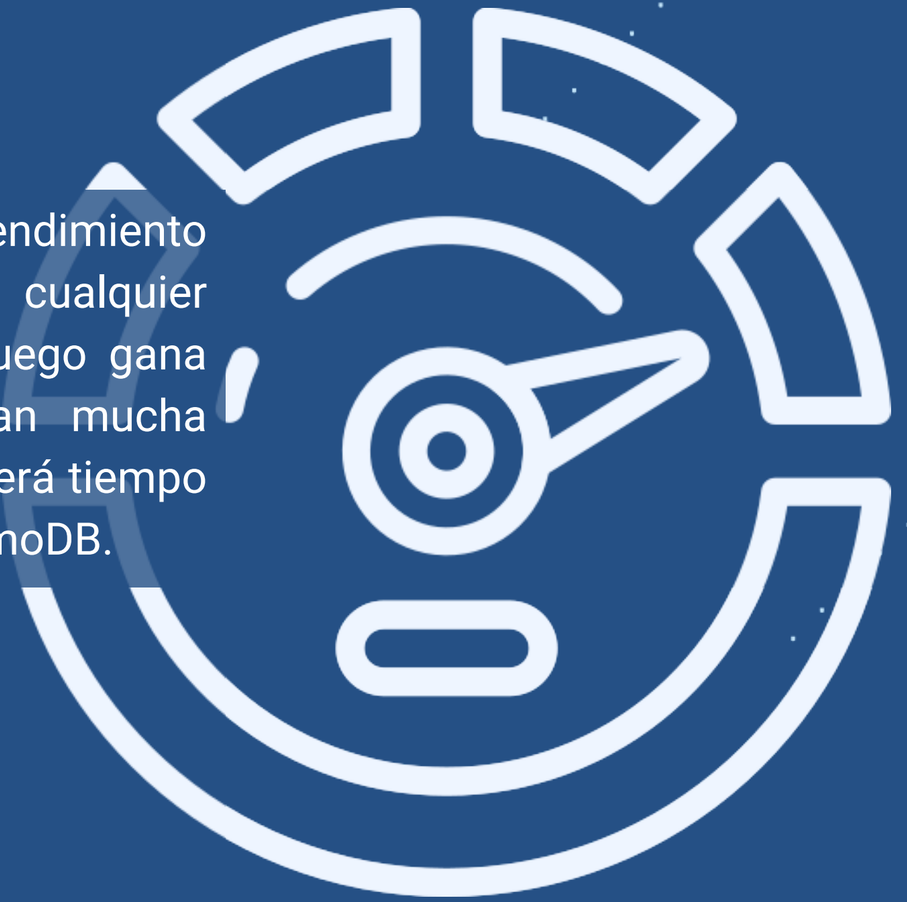
Esta arquitectura podría admitir una aplicación de juegos en línea, en la que la recuperación rápida de sesiones es imprescindible. Los datos de juegos se actualizan continuamente a medida que un jugador acumula objetos, derrota a enemigos, recibe oro, desbloquea niveles y completa logros. Cada evento de sesión debe escribirse en la capa de base de datos para que no se pierda. Los creadores de juegos almacenan el historial de sesiones y otros datos temporales en DynamoDB para lograr una búsqueda rápida por jugador, fecha y hora.



Cada tabla de base de datos de DynamoDB está asociada a una capacidad de rendimiento. Puede especificar 1000 escrituras por segundo, y DynamoDB escala la base de datos en segundo plano. A medida que cambien sus necesidades, puede actualizar la capacidad y Amazon DynamoDB vuelve a asignar recursos según sea necesario. Esta elasticidad ayuda a los desarrolladores de juegos; si su juego se vuelve popular, podría escalar de repente de unos pocos miles de jugadores a millones de ellos. Puede reducir ese número de forma rápida y sencilla si resulta necesario.

COLOMBIA POTENCIA DE LA VIDA

TIC



DynamoDB mantiene un rendimiento predecible de baja latencia a cualquier escala, lo que es crucial si su juego gana millones de clientes que le dan mucha importancia a la latencia. No perderá tiempo ajustando el rendimiento de DynamoDB.

Para ver una arquitectura de juegos similar que incluya a DynamoDB, consulte esta publicación del blog de Big Data de AWS.

TALENTO TECH

AZ | PROYECTOS EDUCATIVOS

UTP Universidad Tecnológica de Pereira

Estos son algunos de los aprendizajes clave de esta lección:

- Las sesiones se utilizan para administrar la autenticación de los usuarios y almacenar sus datos mientras interactúan con la aplicación.
- Puede administrar las sesiones con sesiones rápidas, que es una característica de los balanceadores de carga de Elastic Load Balancing. Las sesiones rápidas dirigen las solicitudes al servidor que administra la sesión del usuario específicamente.
- También puede administrar sesiones conservando los datos de sesión fuera de la instancia del servidor web, por ejemplo, en una caché distribuida o una tabla de DynamoDB.

INICIO