Práctica 5. Servicio de almacenamiento basado en primario/copia

Jorge Aznar López

Jorge Fernández Muñoz

20 de diciembre de 2018

Introducción

En esta práctica se tiene como objetivo presentar una solución de tolerancia a fallos con estado implementando en Elixir un sistema replicación Primario/Copia e implementar el servicio de almacenamiento de Primario/Copia para que sea capaz de tratar las distintas peticiones que realicen los clientes de este servicio, así como ser capaz de mantener los datos almacenados una vez se sufran caídas por parte de los servidores del servicio de almacenamiento que hagan variar el estado del sistema.

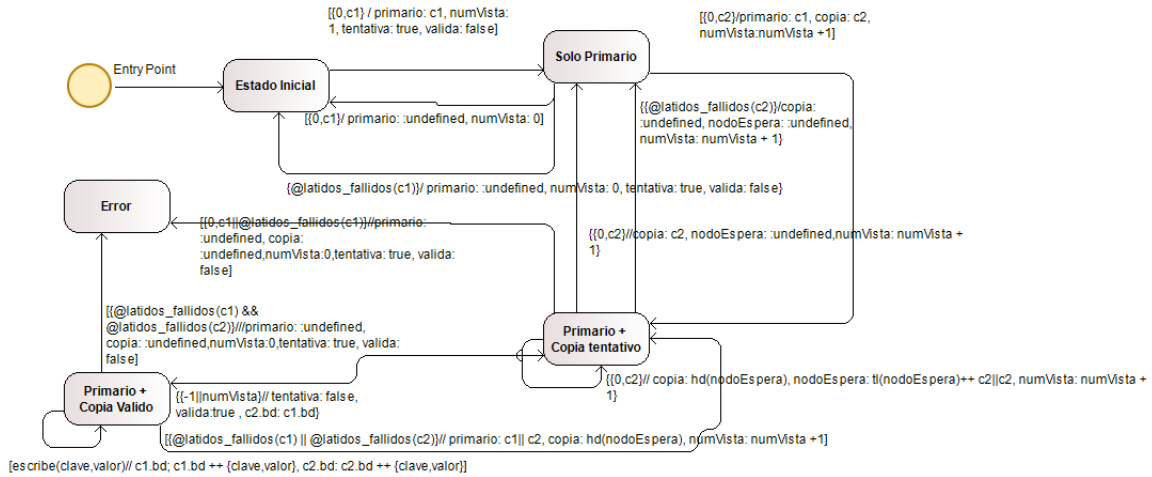
Se ha implementado el módulo del servidor de almacenamiento para que sea capaz de recibir distintas peticiones de lectura y escritura sobre sus datos, además de soportar la caída del nodo primario o copia de este servicio pudiendo garantizar que no se producirá pérdida de datos, en caso de caída simultánea se produciría un fallo crítico del sistema. Asimismo, existen nodos en espera que serían promocionados a nodos copia en caso de caer alguno de los nodos que se encuentran prestando el servicio.

Todos los nodos envían su mensaje de latido al servidor hasta que fallan, y este les responde con la vista tentativa actual, la cual utilizan para conocer su posición dentro del sistema, ya que en caso de ser nodo primario, deberá ser el encargado de recibir y tratar las peticiones de los clientes del servicio y además será el encargado de garantizar que el nodo copia actualice sus datos, añadiendo a estos los datos de las nuevas peticiones, antes de responder a las peticiones, garantizando así que en caso de la caída de uno de estos dos nodos, el sistema podrá seguir en funcionamiento, dado que ambos nodos contendrían la misma información.

Arquitectura del sistema

El sistema de almacenamiento se compone de un servidor gestor de vistas, y de tres nodos servidores del servicio de almacenamiento, así como de clientes de este servicio de almacenamiento. El servidor gestor de vistas despliega un único proceso encargado de recibir los latidos de los distintos servidores del servicio de almacenamiento indicándole así, que siguen activos, por otro lado, estos servidores del servicio de almacenamiento desplegarán dos procesos, uno será el encargado de enviar los latidos al servicio gestor de vistas, recibiendo mensajes de este servicio indicando el estado actual del sistema y la posición que ocupa el servidor del servicio de almacenamiento en el mismo, y otro proceso, que será el encargado de recibir y procesar las peticiones de los clientes del servicio de almacenamiento. Estos clientes del servicio de almacenamiento desplegarán un único proceso que será el encargado de enviar las peticiones.

Estado del sistema

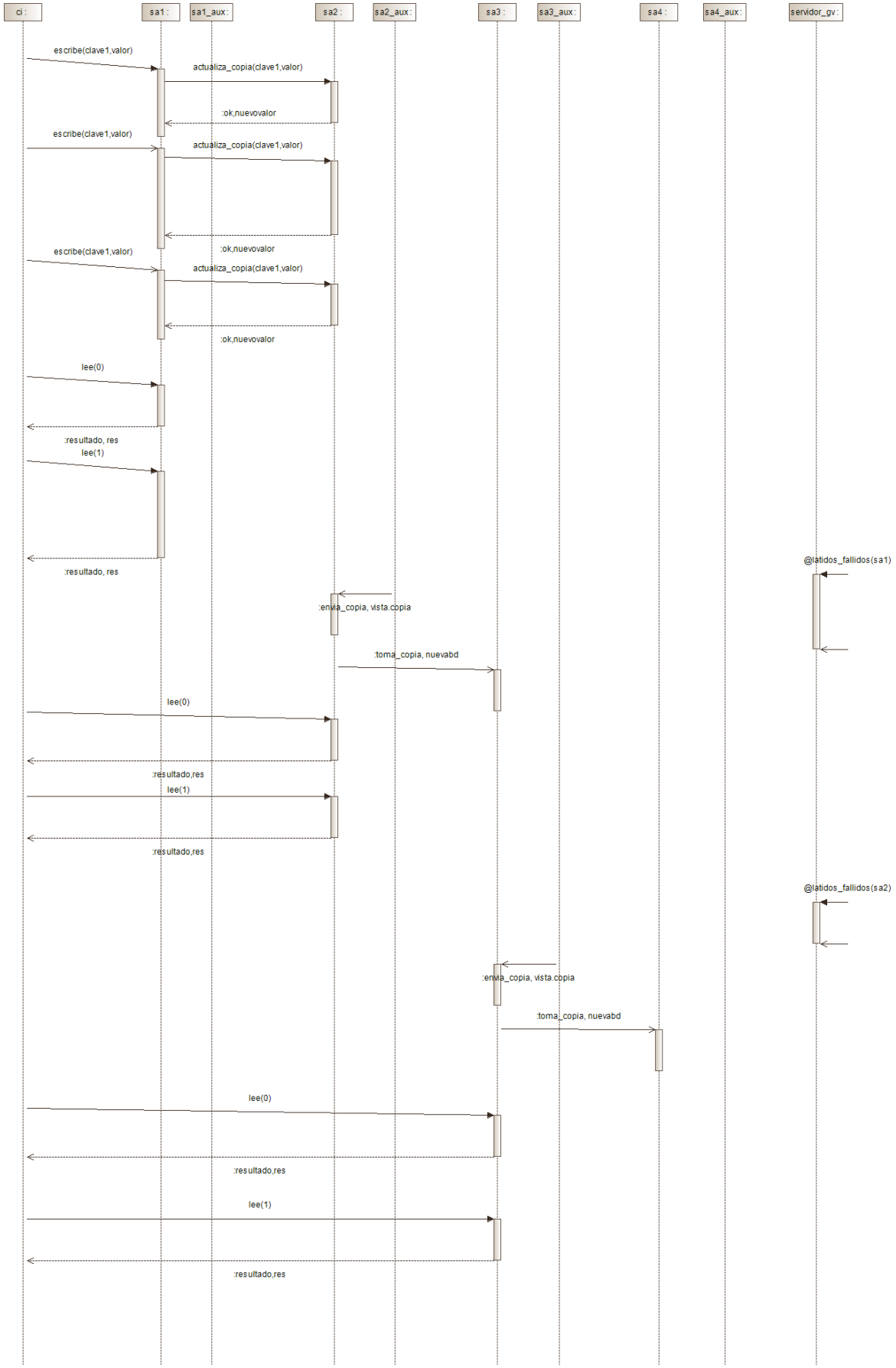


La imagen superior refleja la máquina de estados que representaría el funcionamiento del sistema descrito. Por limpieza se han obviado algunos cambios de estado.

En lo referente al servicio de almacenamiento, destacan dos cambios de estado, el primero el producido una vez nos encontramos en el estado “Primario + Copia tentativo” del cual pasamos al estado “Primario + Copia Valido” una vez el nuevo nodo copia ha copiado la información que contenía el nodo primario, para garantizar la continuidad del sistema en caso de caída de uno de estos nodos.

Por otra parte, el otro cambio de estado importante referente a este servicio de almacenamiento que se ha incluido ocurre en el estado “Primario + Copia Valido”, cuando ya podemos recibir y tratar las peticiones de los clientes, en el caso de que se reciba una petición de escritura, se añadirá a la información ya almacenada.

Procesamiento de las caídas de los servidores del servicio de almacenamiento



Caídas en los servidores del servicio de almacenamiento

Existen dos tipos de caídas posibles en los nodos servidores del sistema de almacenamiento, en ambas existen pérdidas de datos, pero es posible que el nodo rearranque y envíe su latido tras haber caído lo cual supone una varianza entre ambas.

En caso de que se detecte un fallo por caída sin rearranque, el sistema procederá a promocionar, en caso de existir, a los nodos siguientes (nodos copia o en espera), siempre y cuando esta caída no suponga un fallo crítico del sistema (caída de nodo primario sin existir copia válida o caída simultánea primario-copia), y además realizara la copia de la información almacenada en el nodo que continua activo al nuevo nodo promocionado.

Si existe rearranque por parte del nodo caído previamente con el envío de un nuevo latido el sistema será consciente de la posición anterior del nodo caído, realizará la misma promoción de nodos explicada anteriormente y degradará al nodo caído a último nodo en espera.

Validación Experimental

Se ha validado el correcto funcionamiento del sistema con los tres test proporcionados y se ha cumplimentado la validación experimental con el cuarto test especificados que ha sido implementado aparte. Esta validación se ha llevado a cabo en un escenario donde existen tres nodos clientes del sistema de almacenamiento y tres nodos servidores del sistema de almacenamiento, contando el escenario del test 4 con cuatro clientes del sistema de almacenamiento y 4 nodos servidores de este sistema de almacenamiento.

Conclusiones

Se ha corroborado mediante la validación el correcto funcionamiento del sistema en el entorno especificado, es decir, el servicio de almacenamiento es tolerante a fallos no críticos sin que esto suponga la pérdida de datos.