

CC5303 – Sistemas Distribuidos

7.- Replicación

Parte 3

Sebastián Blasco V.

Contenidos

- Replicación
- Modelos de consistencia
 - Centrados en los datos
 - Centrados en el cliente
- Administración de réplicas
- **Protocolos de consistencia**

Protocolos de Consistencia

- Un protocolo de consistencia describe una **implementación** de un modelo específico de consistencia.
- ¿Modelos más importantes y usados?
 - Aquellos en los cuales las operaciones están **globalmente serializadas**
 - Consistencia secuencial
 - Consistencia débil con variables de sincronización
 - Transacciones atómicas.
- Los protocolos pueden ser:
 - Basados en el primario
 - De escritura replicada.

Basados en el primario

- En estos protocolos, cada elemento de datos x del almacén de datos tiene una primaria asociada, la cual es responsable de coordinar las operaciones de escritura sobre x .
- Es posible diferenciar
 - Si la primaria está fija en un **servidor remoto** o
 - *Protocolos de escritura remota*
 - si las operaciones de escritura pueden realizarse **localmente después de trasladar la primaria** al proceso en donde se inició la operación de escritura.
 - *Protocolos de escritura local*

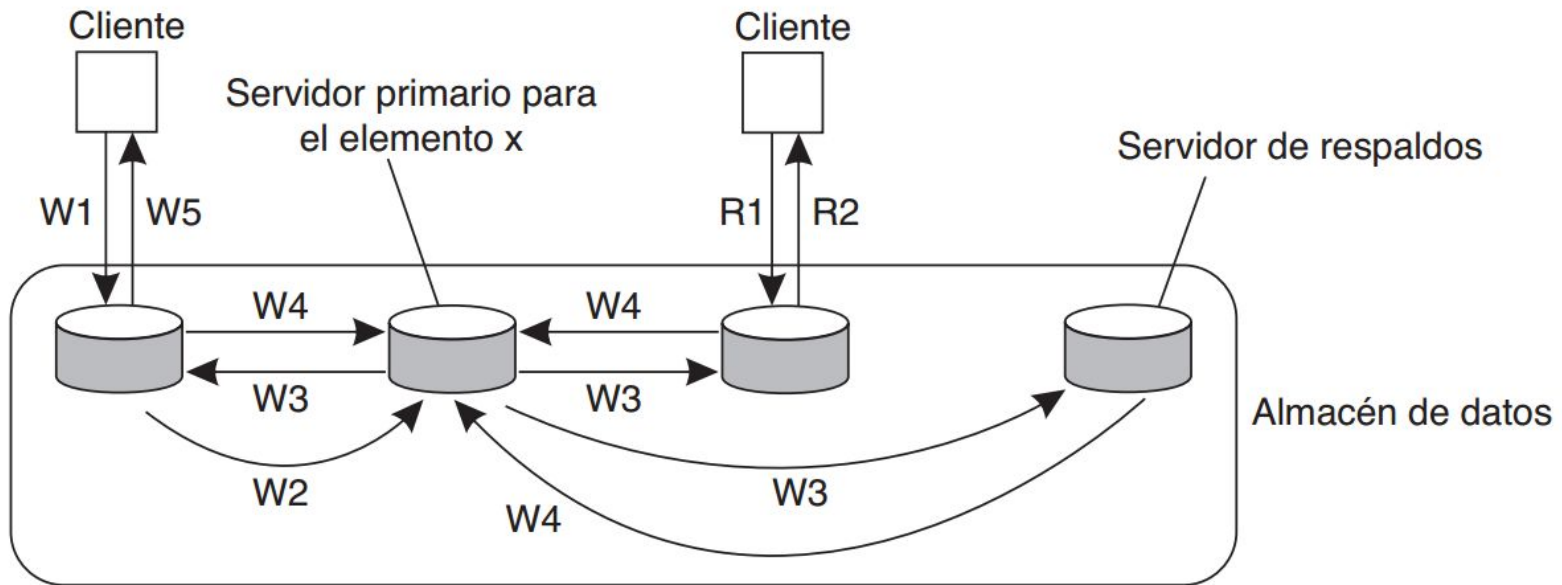
Basados en el primario

Protocolos de escritura remota

- Un proceso que espera para realizar una operación de escritura sobre un elemento de datos x , remite esa operación al servidor primario de x .
- La primaria realiza la **actualización en su copia local de x , y posteriormente remite la actualización** a los servidores de respaldo.
- Cada servidor de respaldo también realiza la actualización, y envía un acuse de recibo de vuelta a la primaria.
- Cuando todos los respaldos han actualizado su copia local, la primaria envía un acuse de vuelta al proceso inicial.

Basados en el primario

Protocolos de escritura remota



- W1. Escribe petición
- W2. Remite la petición a la primaria
- W3. Indica a los respaldos que se actualicen
- W4. Acusa la actualización
- W5. Acusa la escritura finalizada

- R1. Lee la petición
- R2. Respuesta a la lectura

Basados en el primario

Protocolos de escritura remota

- Problema
 - Quien hizo el update (W1) puede demorar mucho en recibir la autorización de continuar (W5), porque **usualmente actualizar es bloqueante**
- ¿Soluciones?:
 - Usar un esquema no bloqueante... lo cual genera problemas de tolerancia a fallas (¿cómo saber qué hizo lo que dijo que haría?)
 - (¿)Cambiar el primario(?)

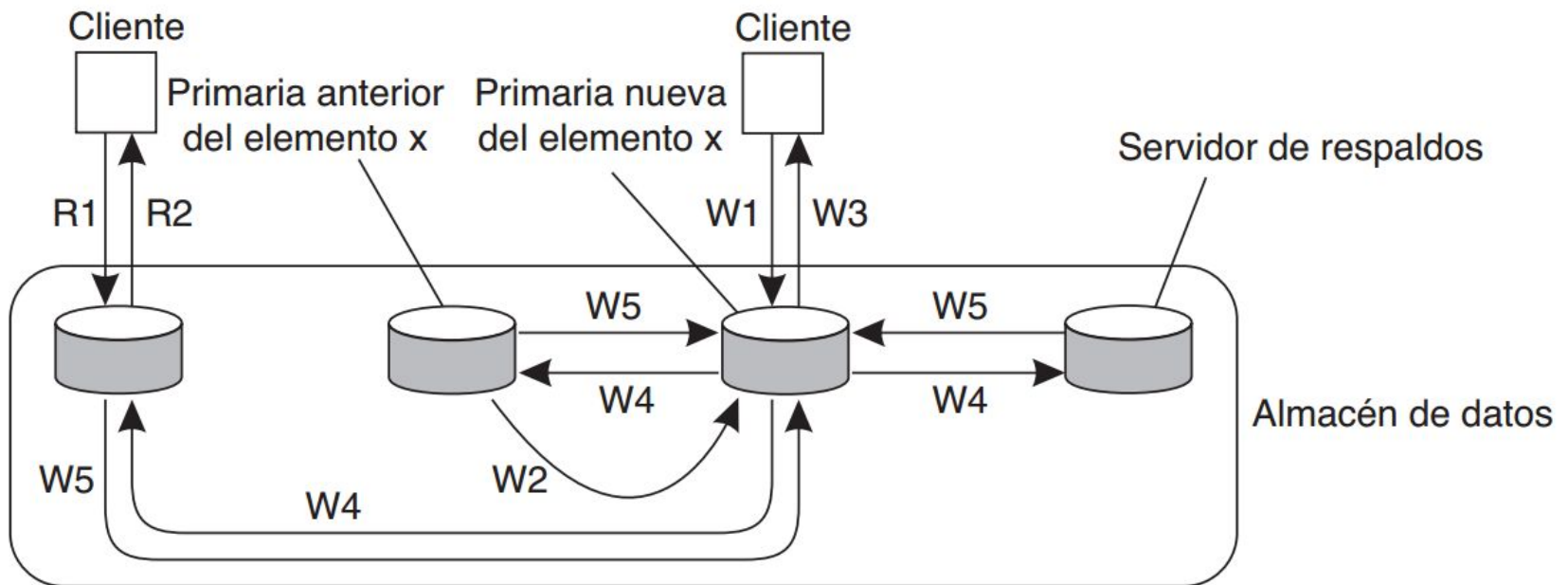
Basados en el primario

Protocolos de escritura local

- Protocolo en el que la copia primaria migra entre procesos que desean realizar una operación de escritura.
 - Siempre que un proceso quiere actualizar el elemento x , éste localiza la copia primaria de x y posteriormente la lleva a su propia ubicación.
- Este protocolo primario de respaldo de escritura local también puede aplicarse a computadoras móviles que son capaces de operar desconectadas.
 - Antes de la desconexión, el cliente se vuelve el primario
 - Todas las operaciones de actualización se llevan a cabo localmente en lo que otros procesos aún pueden realizar operaciones de lectura (pero no actualizaciones).
 - Cuando se conecta de nuevo, las actualizaciones se propagan desde la primaria hacia los respaldos, llevando al almacén de datos de nuevo a un estado consistente.

Basados en el primario

Protocolos de escritura local



W1. Escribe la petición

W2. Lleva el elemento x a la nueva primaria

W3. Acusa la escritura finalizada

W4. Indica a los respaldos que se actualicen

W5. Acusa la actualización

R1. Lee la petición

R2. Respuesta a la lectura

Basados en el primario

Protocolos de escritura local

- Ventaja
 - Varias operaciones sucesivas de escritura pueden realizarse localmente, mientras que los procesos de lectura aún pueden acceder a su copia local.
- Sin embargo,
 - tal mejora sólo puede lograrse si un protocolo de no bloqueo es seguido por las actualizaciones que se propagan a las réplicas después de que la primaria ha terminado de realizar localmente las actualizaciones.

De escritura replicada

- En los protocolos de escritura replicada, las operaciones de escritura pueden ser **ejecutadas en múltiples réplicas** en vez de sólo una.
 - Diferente de el caso de réplicas basadas en primarias.
- Se dividen en
 - Replicación Activa
 - Una operación es llevada a todas las réplicas.
 - Protocolos basados en votación
 - *quorum*

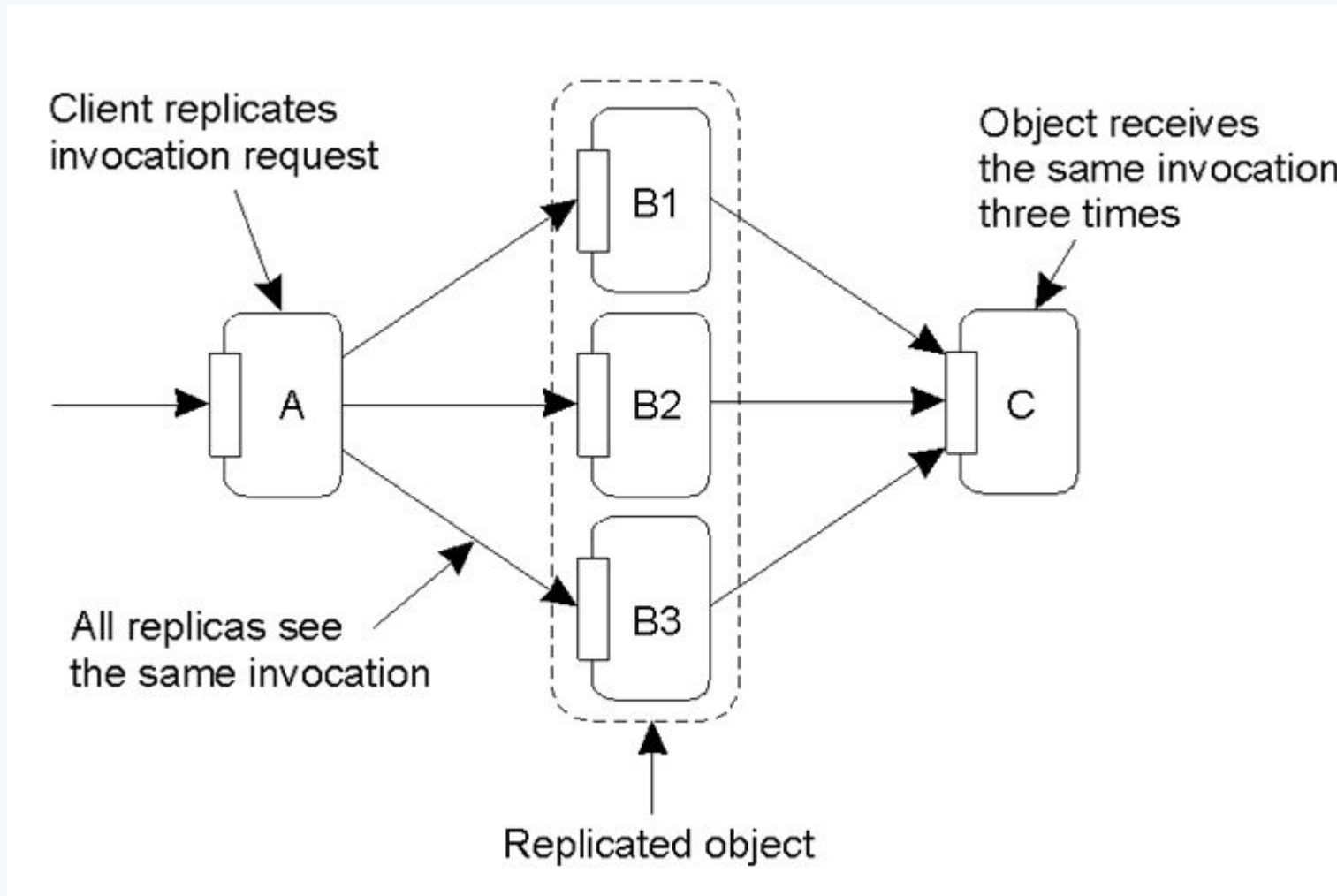
De escritura replicada

Replicación Activa (*Invocaciones Replicadas*)

- Cada réplica tiene asociado un proceso que se encarga de las actualizaciones, **enviando a las otras réplicas la operación que produjo la actualización, no el valor del dato.**
- Beneficios:
 - Reducción de ancho de banda para datos grandes
- Desventaja:
 - Se debe mantener el orden de las operaciones.
 - *¿Qué pasa si cada réplica debía operar además sobre otra réplica?*

De escritura replicada

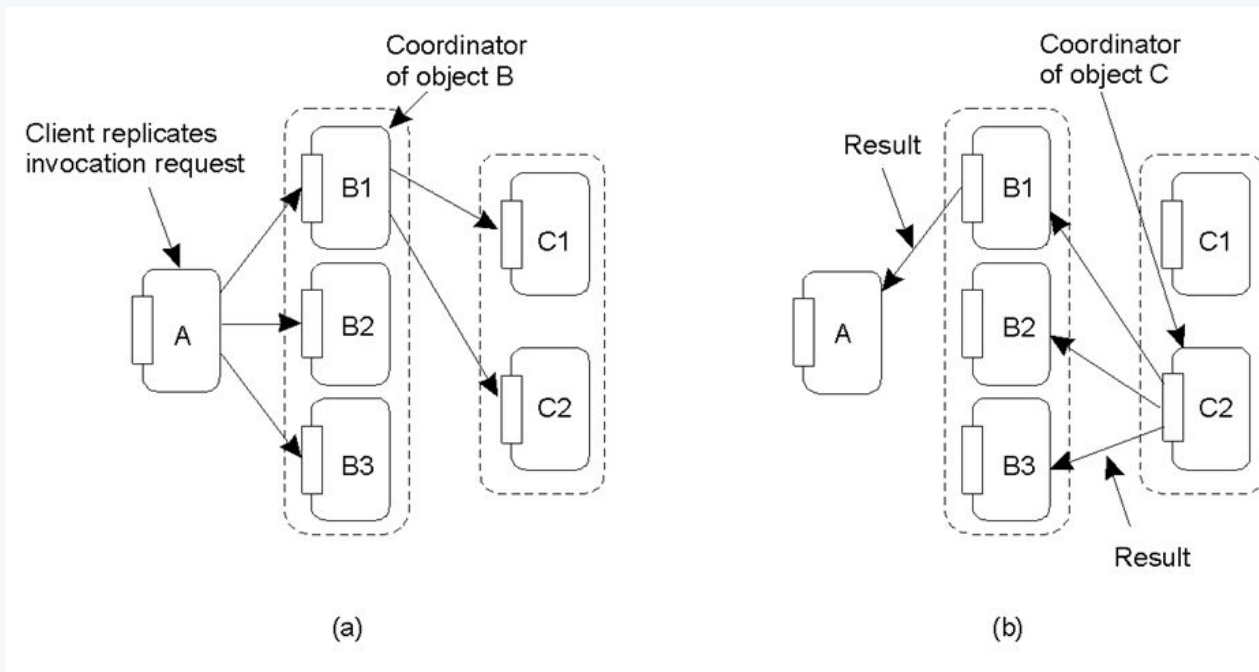
Replicación Activa (*Invocaciones Replicadas*)



De escritura replicada

Replicación Activa (*Invocaciones Replicadas*)

- Solución: **Replication-Awareness**
 - La invocación es asignada a una única réplica mediante la elección de un coordinador. El mismo mecanismo se usa para la respuesta desde las réplicas.



De escritura replicada

Basados en Votación

- La idea básica es requerir a los clientes que soliciten y adquieran el permiso de varios servidores antes de leer o escribir un elemento de datos replicado.
- Para dicho permiso, se impone algún criterio mínimo de votos. Ej:
 - Para **actualizar**: Hacerlo en la mitad más uno de los servidores.
 - Para **leer**: Tener al menos la mitad más uno de los datos en la misma versión.

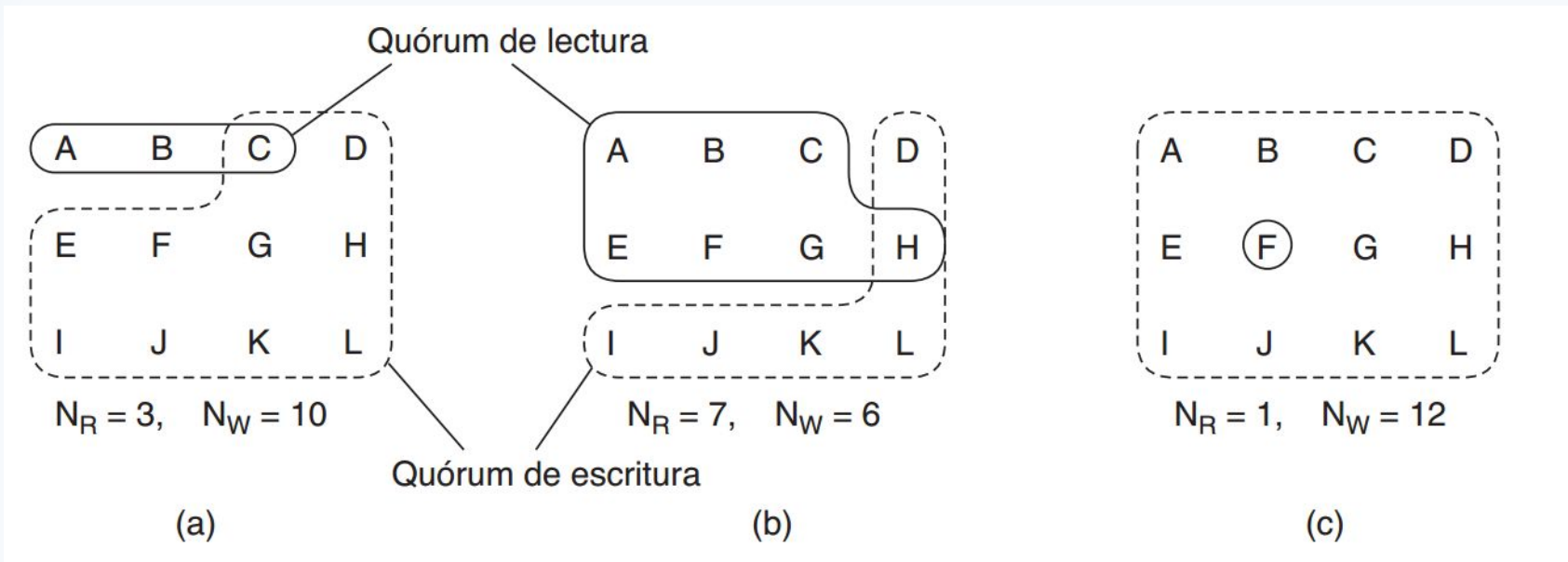
De escritura replicada

Basados en Votación

Sólo después de que el número adecuado de servidores ha acordado participar, un archivo puede leerse o escribirse, bajo las siguientes reglas:

- Sean N réplicas
 - $N_R = \text{Quórum de lectura}$
 - $N_W = \text{Quórum de escritura}$
- Se debe cumplir que:
 - $N_R + N_W > N$
 - Previene conflictos de lectura/escritura
 - $N_W > N/2$
 - Previene conflictos de escritura/escritura

De escritura replicada Basados en Votación



¿Todo Bien?

NO!

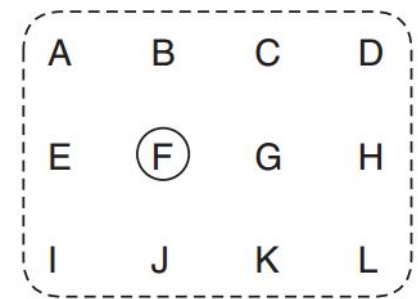
En (b) Si un cliente elige $\{A, B, C, D, F, G\}$ y el otro $\{D, H, I, J, K, L\}$ para escribir, ambas serán aceptadas, pero **habrá conflicto en D**. Esto porque $N_W \leq N/2$

De escritura replicada

Basados en Votación

Caso especial (c)

- $N_R = 1$
- Hace posible leer un archivo replicado al encontrar cualquier copia y usarla.
- El precio pagado por este buen rendimiento de lectura es que las actualizaciones de lectura necesitan adquirir todas las copias.
- Esquema **ROWA**



$$N_R = 1, \quad N_W = 12$$

(c)

Protocolos de Consistencia (Caché)

Extra: El enfoque de consistencia se repite para el caso de las **caché**. Se definen estrategias que definen los posibles protocolos

- No réplica, no migratoria (NRNMB)
 - cliente – servidor.
- No réplica, migratoria (NRMB)
 - Lleva bloques a quien lo usa.
- Réplica, migratoria (RMB)
 - Lleva la réplica más cercana a quien lo usa.
- Réplica, no migratoria (RNMB)
 - Acceso a la réplica más cercana.

Protocolos de Consistencia (Caché)

- Tipos de bloques
 - No usado
 - Read-only
 - Read-owned
 - Escribible
- Escritura invalidante
 - El nodo con la última actualización se encarga de invalidar las otras.
- Escritura actualizada
 - Ordenamiento total de las operaciones de escritura para lograr la consistencia secuencial.
 - Utiliza un secuenciador global.