



INICIO GRABACIÓN



SANJOSÉ
FUNDACIÓN DE EDUCACIÓN SUPERIOR

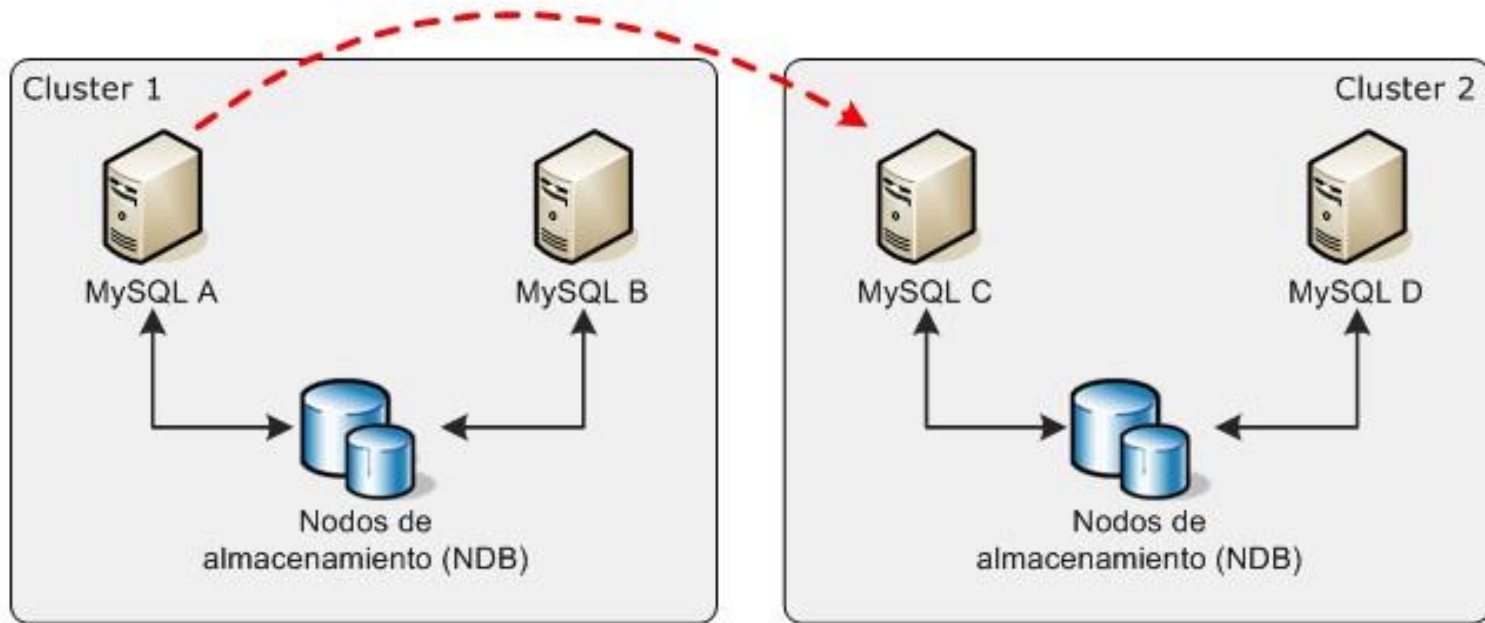
The background features a photograph of several hands of different skin tones cupping a small globe. Overlaid on this are two large, semi-transparent blue circles and a smaller solid blue circle, creating a layered effect.

INTRODUCCION A LA CONSISTENCIA Y REPLICACIÓN

Razones para la replicación



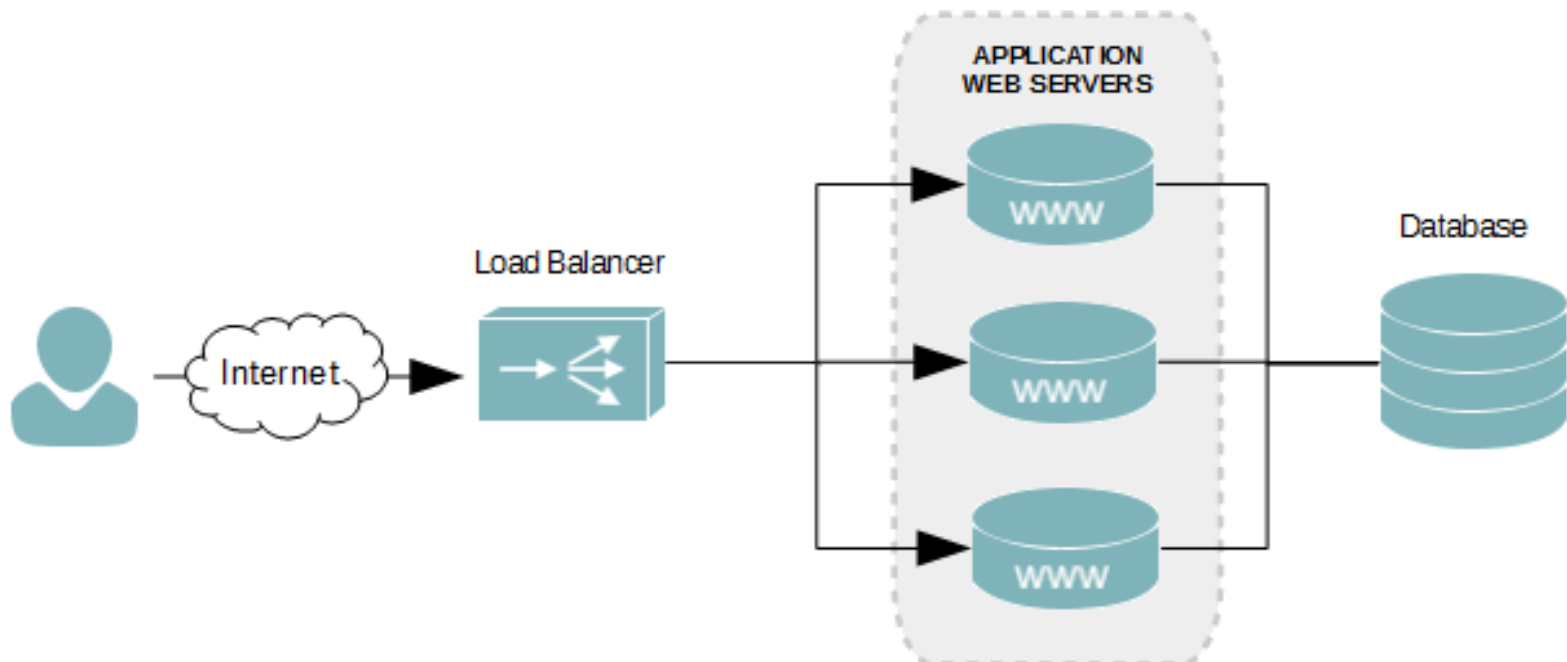
Existen dos razones principales para replicar datos: la confiabilidad y el rendimiento. El problema con la replicación es que tener muchas copias puede provocar problemas de consistencia.



Replicación como técnica de escalamiento



Las cuestiones de escalabilidad generalmente aparecen en forma de problemas de rendimiento. Colocar copias de datos cerca de los procesos que los utilizan puede mejorar el rendimiento mediante la reducción del tiempo de acceso, y resolver así los problemas de escalabilidad.





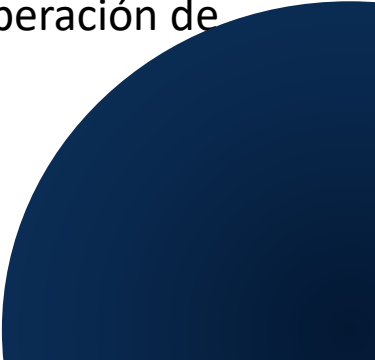
MODELOS DE CONSISTENCIA CENTRADA EN LOS DATOS



Un almacén de datos puede estar físicamente distribuido en varias máquinas. En particular, se asume que todo proceso que puede acceder a datos del almacén tiene una copia local (o en las cercanías) disponible de todo el almacén.

Las operaciones de escritura se propagan hacia las otras copias, Una operación de datos se clasifica como una operación de escritura cuando ésta cambia los datos, de otro modo se clasifica como una operación de lectura.

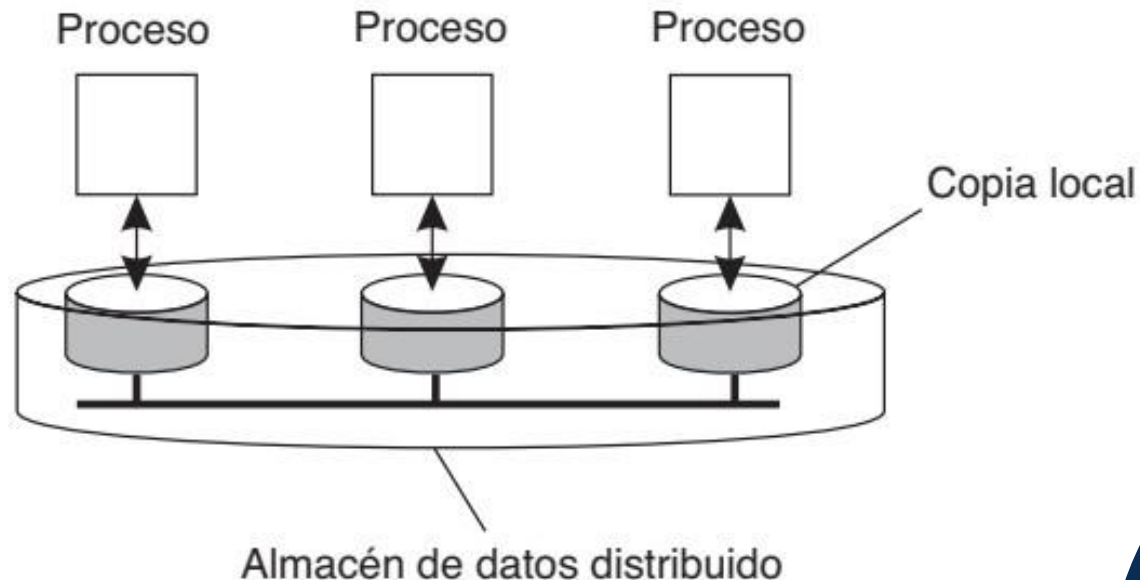
Un modelo de consistencia es básicamente un contrato entre los procesos y el almacén de datos. Este contrato dice que si los procesos aceptan obedecer ciertas reglas, el almacén promete funcionar correctamente. En general, un proceso que realiza una operación de lectura sobre un elemento de datos espera que la operación devuelva un valor que muestre los resultados de la última operación de escritura sobre los datos.



Consistencia continua



Un método general para diferenciar tres ejes independientes para definir inconsistencias, se basa en definir: desviación en valores numéricos entre réplicas, desviación en el deterioro entre réplicas, y desviación con respecto al ordenamiento de operaciones de actualización. Estas desviaciones se conocen como rangos de consistencia continua.



Ordenamiento consistente de operaciones



Todos los algoritmos tienen como base el mismo modelo del sistema. Se supone que cada máquina tiene un cronómetro que ocasiona una interrupción H veces por segundo. Cuando este cronómetro se apaga, el manipulador de interrupciones agrega 1 al reloj de software, el cual da seguimiento al número de marcas (interrupciones) a partir de algún momento pasado acordado.

P1:	W(x)a	
P2:		R(x)NIL R(x)a

Comportamiento de dos procesos operando sobre el mismo elemento de datos. El eje horizontal representa el tiempo.



MODELOS DE CONSISTENCIA CENTRADA EN EL CLIENTE



Este modelo de consistencia está orientado a un conjunto de datos que tienen un bajo número de actualizaciones simultáneas o, cuando estas ocurren, pueden resolverse fácilmente. Generalmente está orientado a operaciones de lectura de datos, por lo que se puede catalogar como un modelo de consistencia bastante débil. Los modelos de consistencias basados en el cliente permiten ver que es posible ocultar muchas inconsistencias de un sistema distribuido de una manera relativamente fácil.

Consistencia momentánea.

Lecturas monotónicas.

Escrituras monotónicas.

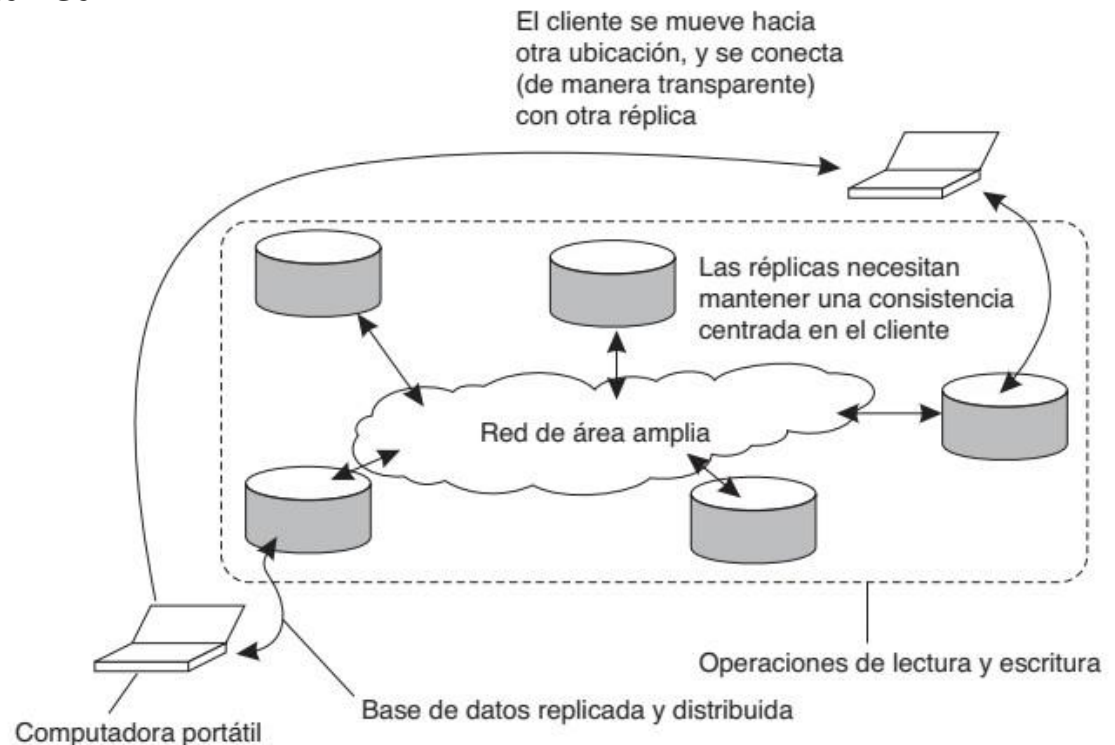
Lea sus escrituras.

Las escrituras siguen a las lecturas.



Consistencia momentánea

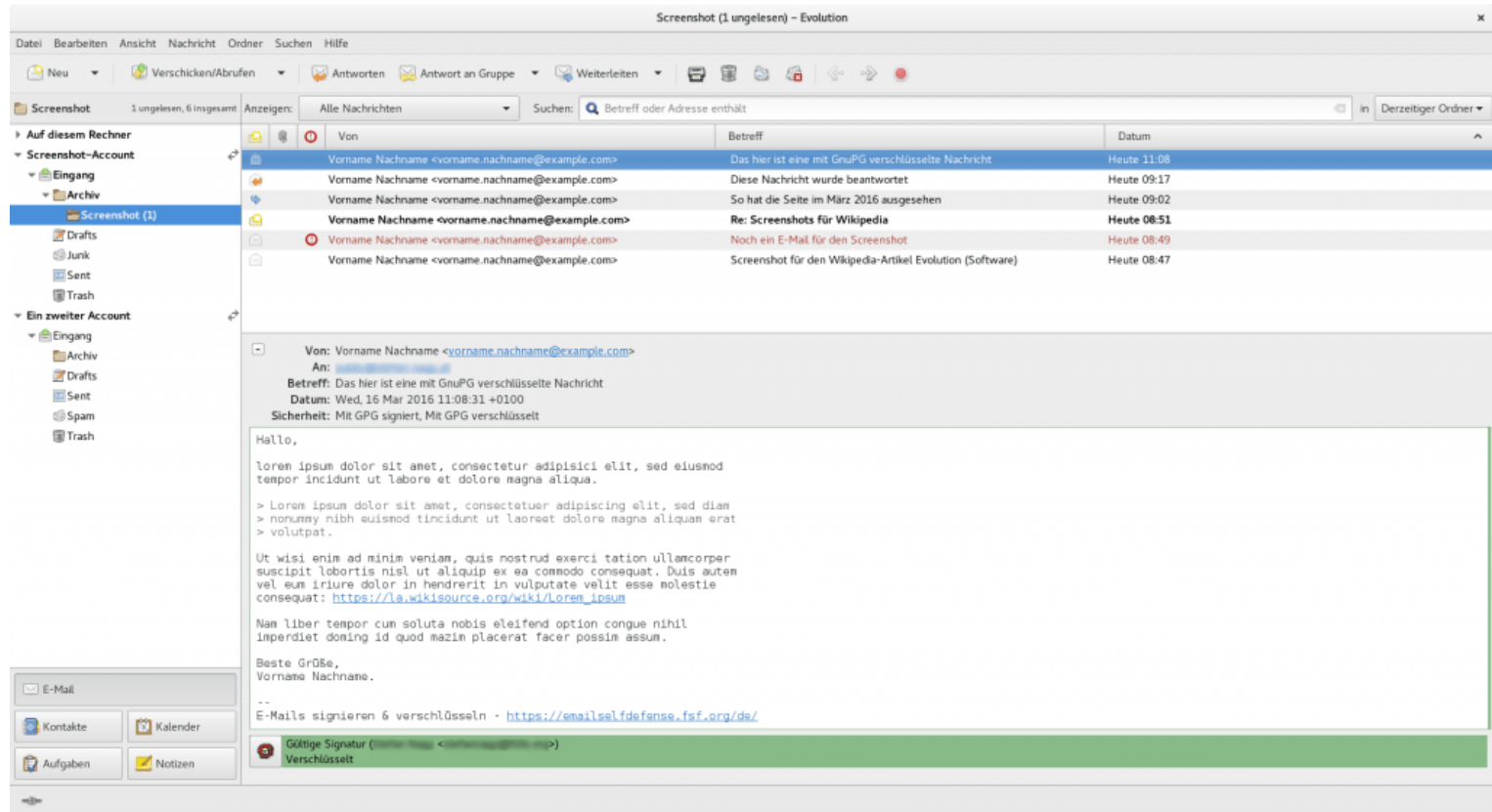
Con frecuencia es aceptable propagar una actualización de modo lento, lo que significa que un proceso de lectura verá una actualización sólo cierto tiempo después de que la actualización ocurrió. En estos casos si no ocurren actualizaciones durante mucho tiempo, todas las réplicas gradualmente se volverán inconsistentes. Esta forma de consistencia se conoce como consistencia momentánea.



Lecturas monotónicas



Se dice que un almacén de datos proporciona consistencia **de lectura monotónica** si se cumple la siguiente condición: *Si un proceso lee el valor de un elemento de datos x , cualquier operación de lectura sucesiva sobre x que haga ese proceso devolverá siempre el mismo valor o un valor más reciente.*



Escrituras monotónicas



En un almacén con consistencia de escritura **monotónica**, se cumple la siguiente condición: *Una operación de escritura hecha por un proceso sobre un elemento x se completa antes que cualquier otra operación sucesiva de escritura sobre x realizada por el mismo proceso.*



Lea sus escrituras

Se dice que un almacén de datos proporciona consistencia **lea sus escrituras** si cumple la siguiente condición: *El efecto de una operación de escritura hecha por un proceso sobre un elemento de datos x siempre será visto por una operación de lectura sucesiva sobre x hecha por el mismo proceso.*



The screenshot shows the 'grupohost' control panel interface. At the top, there is a dark blue header with the 'grupohost' logo on the left, a user profile icon and email 'EJEMPLO@SOPORTE.HOST' in the center, and a 'CIERRE DE SESIÓN' button on the right. Below the header, the main content area has a light gray background. It features a 'Change Password' section with a user icon and a lock icon. The instructions state: 'Escriba una nueva contraseña para esta cuenta.' There are two input fields: 'Nueva contraseña' and 'Confirmar la nueva contraseña', both masked with dots. Below these is a 'Seguridad de la contraseña' section with an information icon. A green progress bar indicates the password is 'Muy segura (100/100)'. To the right of the bar is a 'Generador de contraseñas' button. At the bottom left of the form is a blue 'Guardar' button.

grupohost
control panel

EJEMPLO@SOPORTE.HOST CIERRE DE SESIÓN

Change Password

Escriba una nueva contraseña para esta cuenta.

Nueva contraseña

.....

Confirmar la nueva contraseña

.....

Seguridad de la contraseña ⓘ

Muy segura (100/100)

Generador de contraseñas

Guardar

Las escrituras siguen a las lecturas



Se dice que un almacén de datos proporciona consistencia **las escrituras siguen a las lecturas** si cumple con lo siguiente: *Se garantiza que la operación de escritura de un proceso sobre un elemento de datos x que sigue a una operación de lectura previa sobre x efectuada por el mismo proceso ocurrirá en el mismo o en el más reciente valor de x que se leyó.*





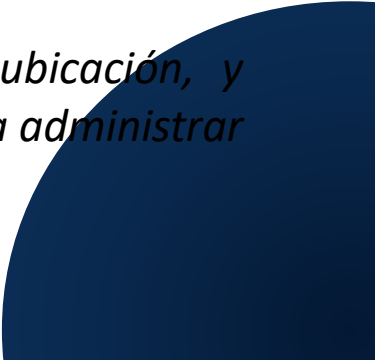
ADMINISTRACIÓN DE RÉPLICAS

Un punto clave para cualquier sistema distribuido que soporta la replicación es decidir dónde, cuándo, y por quién deben ubicarse las réplicas, y posteriormente cuáles mecanismos utilizar para mantener consistentes a dichas réplicas. El problema de ubicación, por sí mismo, debe dividirse en dos subproblemas: el de la ubicación de servidores de réplicas, y el de ubicación de contenido.

La diferencia es sutil pero importante, y con frecuencia las dos cuestiones no están claramente separadas.

La ubicación del servidor de réplicas tiene que ver con encontrar los mejores lugares para colocar un servidor que pueda hospedar (parte de) un almacén de datos. La ubicación de contenido se relaciona con encontrar a los mejores servidores para colocar el contenido. Observe que, generalmente, esto significa que buscamos la ubicación óptima de un solo elemento de datos. Desde luego, antes de que la ubicación de contenido pueda ocurrir, primero tenemos que ubicar los servidores.

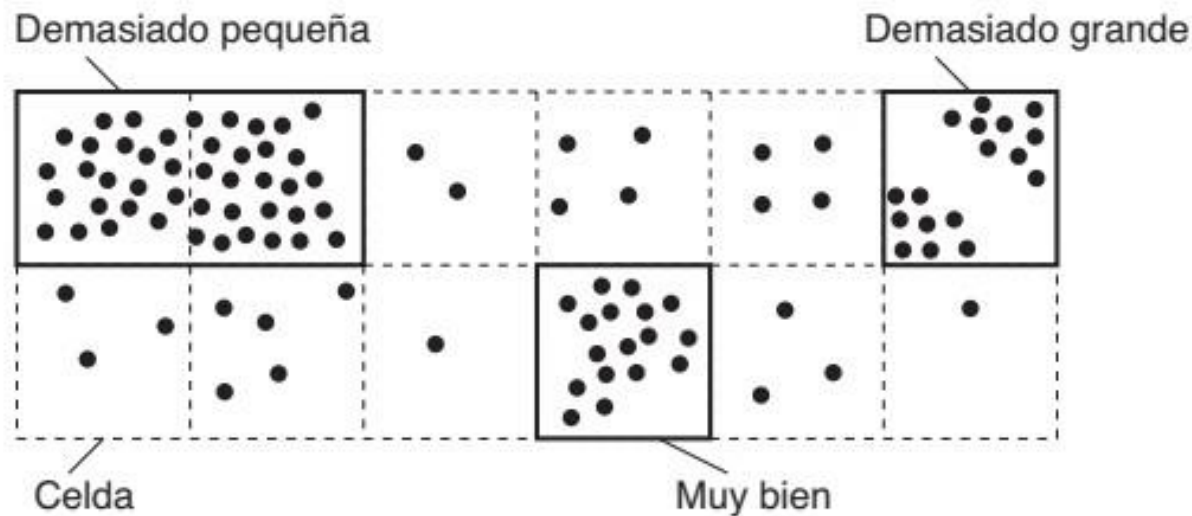
A continuación, daremos un vistazo a estos dos problemas de ubicación, y continuaremos con una explicación sobre los mecanismos básicos para administrar el contenido replicado.



Ubicación del servidor de réplicas



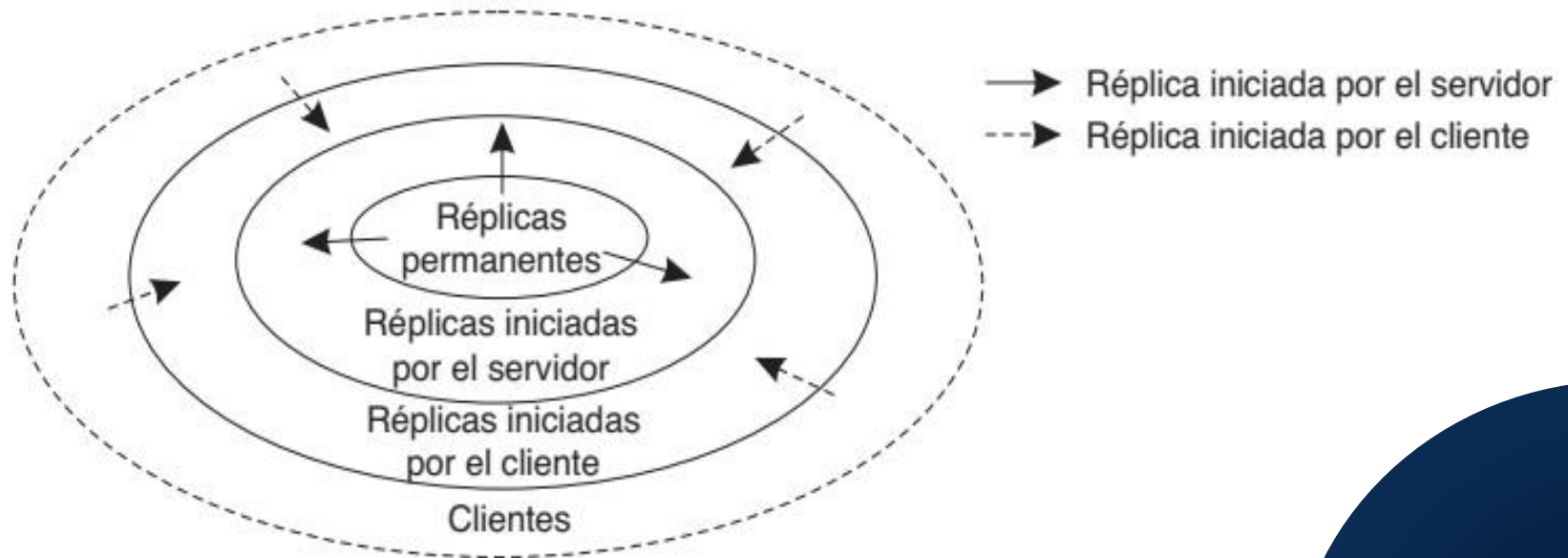
La ubicación de servidores de réplicas no es un problema estudiado intensivamente, por la simple razón de que a menudo se trata de un asunto más administrativo y comercial que un problema de optimización. No obstante, el análisis de las propiedades del cliente y de la red es útil para tomar decisiones informadas.



Ubicación y replicación de contenido



Cuando se trata de replicación y ubicación de contenido, es posible diferenciar tres tipos de réplicas lógicamente organizadas: *Réplicas permanentes*, *Réplicas iniciadas por el servidor* y *Réplicas iniciadas por el cliente*.





Distribución de contenido

La administración de réplicas también trata con la propagación (actualización) de contenido hacia los servidores réplica significativos.

Cuestión	Basado en push	Basado en pull
Estado en el servidor	Lista de réplicas cliente y cachés	Ninguno
Mensajes enviados	Actualiza (y posiblemente trae la actualización después)	Sondea y actualiza
Tiempo de respuesta en el cliente	Inmediato (o busca el tiempo de actualización)	Busca el tiempo de actualización



FUNDACIÓN DE EDUCACIÓN SUPERIOR

SAN JOSÉ

INSTITUCIÓN TECNOLÓGICA

FIN DE
GRABACIÓN