

RC: Actividades Tema 3

ACTIVIDAD 1

Los sistemas de comunicación en grupos tienen sentido para sistemas robustos cuyo objetivo sea mantener una consistencia fuerte. Para ello se utiliza el concepto de vista y se necesita un servicio de pertenencia que “dirija” la transición entre vistas sucesivas. La comunicación es, en principio, asincrónica. El emisor no necesita suspenderse una vez ha iniciado el envío. Sin embargo, muchos servicios de comunicación en grupos utilizan grupos cerrados. En un grupo cerrado las difusiones sólo puede realizarlas un miembro del grupo y tienen siempre como destinatarios a todos los procesos del grupo. Por regla general, aunque la primitiva de difusión no bloquee, el algoritmo que la utiliza suele esperar hasta recibir la difusión para iniciar su siguiente paso.

Razone si se podrán implantar o no servicios altamente escalables (o, simplemente, que admitan altas tasas de envío; por ejemplo, 1000 mens./seg con mensajes de 40 KB cada uno sobre una red local de 1Gbps) con estos tipos de sistemas de comunicación.

No se podrá implantar en un sistema de servicios altamente escalables con este sistema de comunicación de grupos cerrados porque no admite alta disponibilidad dado el solape en el envío de la información. Si tenemos 1000 mensajes de 40kb por segundo, se necesitaría enviar cada uno de esos mensajes a cada uno de los nodos replica. Como cada uno de los mensajes esperan una confirmación de recepción si se tiene un número alto de nodos replica, aumentará el tiempo de respuesta del sistema. Este sistema es útil para sistemas que tienen dos o tres nodos replica.

En los sistemas de computación en la nube no es raro que se utilicen múltiples centros de datos, separados geográficamente, potenciando la escalabilidad de distancia. ¿Son aplicables los servicios de comunicación en grupos en esos entornos? ¿Por qué?

Se podría utilizar porque se utilizarían con pocos nodos y no se saturaría la red al comunicarse el uno con el otro.

ACTIVIDAD 2

Consulte la documentación de un servicio de comunicaciones asíncrono y escalable como ZeroMQ (<http://www.zeromq.org>). Compare sus características con los de los sistemas de comunicación en grupos explicados en este tema 3. ¿Qué ventajas e inconvenientes aporta ZeroMQ en los siguientes aspectos, suponiendo que se usan sus sockets PUB/SUB para realizar las difusiones?

1. Rendimiento (esto es, número de mensajes que podrán llegar a transmitirse a un determinado conjunto de procesos por unidad de tiempo).

ZeroMQ está orientado a la escalabilidad, lo que significa que un numero muy grande de datos se puede enviar de manera más rápida a multiples nodos. Si se envían por ejemplo un millón de mensajes por segundo a cientos de nodos, PUB/SUB de ZeroMQ es una buena solución. [<http://zguide.zeromq.org/php:chapter5>]

2. Garantías de entrega.

PUB/SUB no se garantiza la entrega del mensaje a todos los nodos. Si un instante t un nodo V se cae, luego en el instante $t+1$ se envía un mensaje M y el nodo V se recupera en el instante $t+2$, el mensaje M no será recibido por el nodo V . [<http://zguide.zeromq.org/php:chapter5>]

3. Garantías de orden y consistencia.

Si es un solo Publicador se garantiza el orden de llegada de los mensaje. También se garantiza la consistencia.

Si hay más de un Publicador no se garantiza el orden de llegada de los mensajes. Tampoco se tendría consistencia porque el orden de llegada de los mensajes afectaría el funcionamiento del nodo. [<http://zguide.zeromq.org/php:chapter5>]

4. Sincronía / asincronía en la gestión de la transmisión.

La comunicación de PUB/SUB es asíncrona. Si un nodo publicador ya se ha iniciado y luego se ha iniciado un nodo suscriptor, este no recibiría los mensajes que ya fueron publicados por el nodo publicador. [<http://learning-0mq-with-pyzeromq.readthedocs.org/en/latest/pyzeromq/patterns/pubsub.html>]

5. Detección y gestión de fallos en los procesos del grupo.

En las conexiones de PUB/SUB, si un cliente se desconecta el servidor no se percató de ello. Eso es posible dado que el servidor no recibe ninguna confirmación por parte de los clientes de la recepción del mensaje.

Si el servidor es el que se desconecta, los clientes no tienen por qué percatarse si no se realiza ninguna prevención por parte de los clientes. Si el servidor envía un mensaje de que está vivo periódicamente, al pasar un tiempo si los clientes no reciben ese mensaje pueden sospechar ya que el servidor ha caído. [<http://zguide.zeromq.org/php:chapter4>]

En clase se comentó que ZeroMQ puede tener internamente un sistema de monitorización para detectar la caída de una conexión para evitar realizar envíos innecesarios, pero a la vista del uso por parte del cliente eso no es visible.

ACTIVIDAD 3

Las propiedades 3, 4 y 5 enunciadas en la sección 4.1 (Propiedades de seguridad) proporcionan tres niveles distintos de sincronización en las entregas de los mensajes en caso de que haya cambios de vista en un determinado grupo (debido al fallo o recuperación de uno de los nodos del grupo). Responda las siguientes cuestiones sobre esas propiedades, justificando sus respuestas.

1. ¿Se puede afirmar que la sincronía virtual (Propiedad 5) se cumplirá cuando se cumpla cualquiera de las otras dos propiedades? ¿Por qué?

Sí. Porque si tenemos dos procesos P y Q que instalan la misma vista V sobre la misma vista previa V', entonces cualquier mensaje entregado por P en V' fue también entregado por Q en V'.

Como la propiedad 3 dice que si un proceso P entrega un mensaje m en la vista V y algún proceso Q envió m en la vista V' entonces $V=V'$. Esto está implícito en la definición de la propiedad 5, cuando se explica como se relacionan P y Q con V y V'.

Luego como la propiedad 4 dice que si dos procesos P y Q entregan un mensaje en m, ambos lo entregan en la misma vista. En la propiedad 5 eso se explica al final cuando un mensaje P se envía en V', también será entregado a Q en V'.

2. ¿Qué exigencias adicionales implica la propiedad 3 (entrega en la vista de envío) cuando se compara con la propiedad 4 (entrega en la misma vista)?

Antes de realizar un envío a una vista diferente se tiene que haber acabado el envío inicial o se tienen que descartar los mensajes. Mientras que la 4ª propiedad solo admite comunicación en la misma vista.

3. ¿Tiene sentido la propiedad 3 cuando se asume un modelo de fallos que admita la recuperación de los procesos caídos?

No. No se puede porque no estaba en la vista cuando se envió los mensajes. Por lo tanto no se puede recuperar los mensajes que se enviaron en la vista porque ahora están en una vista diferente.

La propiedad 3 sólo admite que el mensaje m esté en la vista en la que se envió.

4. ¿Tiene sentido la propiedad 5 en un sistema elástico de computación en la nube? ¿Por qué?

No. El coste sería excesivo y la pertenencia al grupo no es tan estricto. La sincronía virtual no permitiría al sistema escalar de forma fácil. La sincronía virtual incluye consenso lo que lo hace muy costoso.

ACTIVIDAD 4

Justifique cuántas rondas de intercambio de mensajes se necesitarán para implantar los siguientes tipos de difusión fiable ordenada:

1. Difusión fiable FIFO.

Difusión FIFO: mediante contadores FBCAST. Cada mensaje tiene un identificador de secuencia asignado. Si el mensaje se recibe tiene el indicador correcto, entonces el mensaje se entrega de inmediato. En cambio si el mensaje que se recibe tiene un identificador que no corresponde al mensaje, entonces es encolado y no será entregado hasta que se haya recibido el mensaje previo. En este caso se necesitarían realizarán M envíos para los M mensajes.

2. Difusión fiable causal.

Difusión causal: mediante relojes vectoriales CBCAST. Cada mensaje tiene un identificador asignado. Si el mensaje se recibe tiene el indicador correcto, entonces el mensaje se entrega de inmediato. En cambio si el mensaje que se recibe tiene un identificador que no corresponde al orden de los mensajes, entonces es encolado y no será entregado hasta que se haya recibido el mensaje previo. En este caso se necesitarían M envíos para los M mensajes.

3. Difusión fiable de orden total.

Difusión fiable de orden total: Mediante gestión de vistas donde un proceso primero se conecta a la vista donde quiere enviar el mensaje y posteriormente envía el mensaje. En este caso se necesitarían M envíos + 1 por M mensajes.

ACTIVIDAD 5

Busque referencias bibliográficas que describan cada una de ellas algún algoritmo de difusión de las dos últimas clases mencionadas en la actividad anterior (aparte de las citadas en la presentación del tema). Para cada una de ellas: proporcione la referencia (citando autor, título, conferencia o revista en la que fue publicada, año, páginas y número o volumen, en caso de que fuese una revista), resuma brevemente el algoritmo utilizado y justifique el coste (en número de rondas de comunicación utilizadas) de la solución presentada.

1. Difusión fiable causal.

“Causally ordering group communication protocol” por Nakamura, A., Tachikawa, T., & Takizawa, M.

El artículo se explica como consiguen ordenar los mensajes de forma casual mediante una secuencia que se le asigna a cada mensaje.

2. Difusión fiable de orden total.

“Non-blocking message total ordering protocol” por Wang, Y., & Wang, J.

En el artículo se explica como han desarrollado un algoritmo de difusión de orden que no se bloquea ante un cambio de vista.