

PRA1. Sistemas distribuidos.

Kepa Sarasola Bengoetxea

1. Phase 1: Theoretical exercise of TSAE protocol

1.1 TSAE protocol exercise (no purged log)

Situación inicial:

Summary A= A3, B3, C2
Summary B= A1, B4, C2
Summary C= A2, B3, C4

Log A= A1, A2, A3, B1, B2, B3, C1, C2
Log B= A1, B1, B2, B3, B4, C1, C2
Log C= A1, A2, B1, B2, B3, C1, C2, C3, C4

Ejercicio 1.1.1

Time	Summary	Log	Operation
T ₀	Summary A= A3, B3, C2 Summary B= A1, B4, C2 Summary C= A2, B3, C4	Log A= A1, A2, A3, B1, B2, B3, C1, C2 Log B= A1, B1, B2, B3, B4, C1, C2 Log C= A1, A2, B1, B2, B3, C1, C2, C3, C4	Initial state
T ₁	Summary A= A4 , B3, C2 Summary B= A1, B4, C2 Summary C= A2, B3, C4	Log A= A1, A2, A3, A4 , B1, B2, B3, C1, C2 Log B= A1, B1, B2, B3, B4, C1, C2 Log C= A1, A2, B1, B2, B3, C1, C2, C3, C4	Host A executes operation A4
T ₂	Summary A= A4, B3, C2 Summary B= A2 , B4, C4 Summary C= A2, B4 , C4	Log A= A1, A2, A3, A4, B1, B2, B3, C1, C2 Log B= A1, A2 , B1, B2, B3, B4, C1, C2, C3 , C4 Log C= A1, A2, B1, B2, B3, B4 , C1, C2, C3, C4	Host B does an anti-entropy session with host C
T ₃	Summary A= A4, B3, C2 Summary B= A2, B4, C4 Summary C= A2, B4, C5	Log A= A1, A2, A3, A4, B1, B2, B3, C1, C2 Log B= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5	Host C executes operations C5
T ₄	Summary A= A4, B4 , C5 Summary B= A2, B4, C4 Summary C= A4 , B4, C5	Log A= A1, A2, A3, A4, B1, B2, B3, B4 , C1, C2, C3 , C4 , C5 Log B= A1, A2, A3 , A4 , B1, B2, B3, B4, C1, C2, C3, C4 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5	Host A does an anti-entropy session with host C
T ₅	Summary A= A5, B4, C5 Summary B= A2, B4, C4 Summary C= A4, B4, C5	Log A= A1, A2, A3, A4, A5 , B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, B1, B2, B3, B4, C1, C2, C3, C4 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5	Host A executes operation A5
T ₆	Summary A= A5, B4, C5 Summary B= A2, B5 , C4 Summary C= A4, B4, C5	Log A= A1, A2, A3, A4, A5 , B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, B1, B2, B3, B4, B5 , C1, C2, C3, C4 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5	Host B executes operations B5
T ₇	Summary A= A5, B5 , C5 Summary B= A5 , B5, C5 Summary C= A4, B4, C5	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5 , B1, B2, B3, B4, B5, C1, C2, C3, C4, C5 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5	Host A does an anti-entropy session with host B

T ₈	Summary A= A5, B5, C5 Summary B= A5, B5, C5 Summary C= A4, B4, C6	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, C1, C2, C3, C4, C5 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5, C6	Host C executes operations C6
T ₉	Summary A= A5, B5, C5 Summary B= A5, B8 , C5 Summary C= A4, B4, C6	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8 , C1, C2, C3, C4, C5 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5, C6	Host B executes operation B6, B7, B8
T ₁₀	Summary A= A5, B5, C5 Summary B= A5, B8, C5 Summary C= A4, B4, C8	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5, C6, C7, C8	Host C executes operations C7, C8
T ₁₁	Summary A= A6 , B5, C5 Summary B= A5, B8, C5 Summary C= A4, B4, C8	Log A= A1, A2, A3, A4, A5, A6 , B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5 Log C= A1, A2, B1, B2, B3, B4, C1, C2, C3, C4, C5, C6, C7, C8	Host A executes operations A6
T ₁₂	Summary A= A6, B5, C5 Summary B= A5, B8, C8 Summary C= A5, B8 , C8	Log A= A1, A2, A3, A4, A5, A6, B1, B2, B3, B4, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8 Log C= A1, A2, A3, A4, A5 , B1, B2, B3, B4, B5, B6, B7, B8 , C1, C2, C3, C4, C5, C6, C7, C8	Host B does an anti-entropy session with host C
T ₁₃	Summary A= A6, B8, C8 Summary B= A5, B8, C8 Summary C= A6 , B8, C8	Log A= A1, A2, A3, A4, A5, A6, B1, B2, B3, B4, B5, B6, B7, B8 , C1, C2, C3, C4, C5, C6, C7, C8 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8 Log C= A1, A2, A3, A4, A5, A6 , B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8	Host C does an anti-entropy session with host A
T ₁₄	Summary A= A7 , B8, C8 Summary B= A5, B8, C8 Summary C= A6, B8, C8	Log A= A1, A2, A3, A4, A5, A6, A7 , B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8 Log B= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8 Log C= A1, A2, A3, A4, A5, A6, B1, B2, B3, B4, B5, B6, B7, B8, C1, C2, C3, C4, C5, C6, C7, C8	Host A executes operation A7

El estado final NO ES CONSISTENTE, ya que todos los host no han recibido las mismas operaciones. Para obtener una situación de consistencia, deberían de realizarse las siguientes operaciones:

- Host A necesita hacer una sesión de anti-entropía con el Host B
- Host A necesita hacer una sesión de anti-entropía con el Host C

Ejercicio 1.1.2

Time	Summary	Log	Operation
T ₀	Summary A= A3, B3, C2 Summary B= A1, B4, C2 Summary C= A2, B3, C4	Log A= A1, A2, A3, B1, B2, B3, C1, C2 Log B= A1, B1, B2, B3, B4, C1, C2 Log C= A1, A2, B1, B2, B3, C1, C2, C3, C4	Initial state
T ₁	Summary A= A4 , B3, C2 Summary B= A1, B4, C2 Summary C= A2, B3, C4	Log A= A1, A2, A3, A4 , B1, B2, B3, C1, C2 Log B= A1, B1, B2, B3, B4, C1, C2 Log C= A1, A2, B1, B2, B3, C1, C2, C3, C4	Host A executes operation A4
T ₂	Summary A= A4, B3, C4 Summary B= A1, B4, C2 Summary C= A4 , B3, C4	Log A= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4 Log B= A1, B1, B2, B3, B4, C1, C2 Log C= A1, A2, A3, A4 , B1, B2, B3, C1, C2, C3, C4	Host A does an anti-entropy session with host C
T ₃	Summary A= A4, B3, C4 Summary B= A1, B6 , C2 Summary C= A4, B3, C4	Log A= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4 Log B= A1, B1, B2, B3, B4, B5, B6 , C1, C2 Log C= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4	Host B executes operation B5, B6
T ₄	Summary A= A4, B6 , C4 Summary B= A4 , B6, C4 Summary C= A4, B3, C4	Log A= A1, A2, A3, A4, B1, B2, B3, B4, B5, B6 , C1, C2, C3, C4 Log B= A1, A2, A3, A4 , B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log C= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4	Host A does an anti-entropy session with host B
T ₅	Summary A= A4, B6, C4 Summary B= A4, B6, C4 Summary C= A4, B3, C5	Log A= A1, A2, A3, A4, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log B= A1, A2, A3, A4, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log C= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4, C5	Host C executes operation C5
T ₆	Summary A= A5 , B6, C4 Summary B= A4, B6, C4 Summary C= A4, B3, C5	Log A= A1, A2, A3, A4, A5 , B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log B= A1, A2, A3, A4, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log C= A1, A2, A3, A4, B1, B2, B3, C1, C2, C3, C4, C5	Host A executes operation A5
T ₇	Summary A= A5, B6, C5 Summary B= A4, B6, C4 Summary C= A5, B6 , C5	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4 Log C= A1, A2, A3, A4, A5 , B1, B2, B3, B4, B5, B6 , C1, C2, C3, C4, C5	Host C does an anti-entropy session with host A
T ₈	Summary A= A5, B6, C5 Summary B= A5 , B6, C5 Summary C= A5, B6, C5	Log A= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4, C5 Log B= A1, A2, A3, A4, A5 , B1, B2, B3, B4, B5, B6, C1, C2, C3, C4, C5 Log C= A1, A2, A3, A4, A5, B1, B2, B3, B4, B5, B6, C1, C2, C3, C4, C5	Host A does an anti-entropy session with host B

El estado final ES CONSISTENTE, ya que todos los host han recibido las mismas operaciones.

1.2 TSAE protocol exercise (purged log)

AckSummary A

A	A	B	C	D	E
	5	2	1	2	3
B	A	B	C	D	E
	4	3	4	2	4
C	A	B	C	D	E
	2	2	5	2	2
D	A	B	C	D	E
	3	3	4	3	4
E	A	B	C	D	E
	4	3	1	1	4

AckSummary B

A	A	B	C	D	E
	4	2	3	4	5
B	A	B	C	D	E
	2	4	3	2	5
C	A	B	C	D	E
	3	2	3	3	5
D	A	B	C	D	E
	3	1	3	4	5
E	A	B	C	D	E
	4	1	3	4	5

a) Which is the content of Log in host A?

Log A : A3, A4, A5, B3, C4, C5, D2, D3, E3, E4

b) Which is the content of Log in host B?

Log B: A3, A4, B2, B3, B4, D3, D4

D and E do an anti-entropy session. During the session both know who is each other (this is different from the algorithm in the Golding thesis).

Summary A = A5, B3, C5, D3, E4

Summary B = A4, B4, C3, D4, E5

c) Which operations are exchanged during the anti-entropy session?

Operaciones enviadas por A = A5, C5

Operaciones enviadas por B = B4, D4, E5

Despues de la sesión de anti-entropía:

Summary A = A5, B4, C5, D4, E5

Summary B = A5, B4, C5, D4, E5

d) Which AckSummary and log have each host after ending the session?

AckSummary A , B = A: {A5, B4, C5, D4, E5}

B: {A5, B4, C5, D4, E5}

C: {A3, B2, C5, D3, E5}

D: {A3, B3, C4, D4, E5}

E: {A4, B3, C3, D4, E5}

Log A, B = A4, A5, B3, B4, C4, C5, D4

1.3 TSAE protocol exercise

Question 1: In TSAE, why is it needed to purge the log periodically? Explain with your own words the question.

Tal y como se indica en el documento TSAE-Golding-1992-Phd, el log de mensajes debe de purgarse periódicamente ya que si no este log crecerá sin ningún limite. El documento nos indica que incluso en sistemas en los que no existe un log, como por ejemplo en aplicaciones que trabajan directamente sobre una base de datos de aplicaciones, los certificados de defunción también deben purgarse.

Existe una condición necesaria para que un mensaje pueda ser borrado del log de mensajes y es que todos los implicados en este lo hayan recibido, esto se puede certificar mediante el vector de marcas de tiempo de acuse de recibo, ya que cuando el mensaje es anterior a todos los eventos del vector de marcas de tiempo de acuse de recibo se puede eliminarse con seguridad.

El vector de acuse de recibo, se actualiza durante cada sesión de entropía llevada a cabo con éxito. En cada una de estas sesiones, se añade al vector de acuse de recibo la información sobre los mensajes que se han recibido y se fija un vector de tiempos para él. Al igual que con los mensajes, este acuse de recibo se propagará a todos los sistemas participantes, dándose por enterado cada uno de ellos de que ha sido recibido y reconocido por todas las partes.

Una vez en este punto, en el que todos los participantes en la comunicación se han dado por enterado y han reconocido los diferentes acuses de recibo, el mensaje habrá sido reconocido por todas las partes y por lo tanto puede ser eliminado. En un escenario de éxito todas las partes reconocerán los diferentes mensajes generados en la comunicación por lo que todos los mensajes podrán ser eliminados.

Question 2: Describe what are the main fields of a TSAE data structure.

Una estructura de datos TSAE, debe mantener a su vez tres estructuras de datos:

- Un registro de mensajes
- Dos vectores de marcas de tiempo

Estas estructuras de datos deben de mantenerse en un almacenamiento estable, para que no se pierdan si el host sufre un problema y no puede responder a las peticiones.

El registro de mensajes

Contiene los mensajes que ha recibido un host, cuando se recibe un mensaje, a este se le añade una marca de tiempo y se añade al registro, este mensaje se elimina del registro cuando se comprueba que todos los demás host también lo han recibido.

Los mensajes se entregan al host una vez se han registrado, ya que es el registro de mensajes quien se encarga de notificar al host el mensaje. El registro de mensajes incluye funciones para recuperar los mensajes enviados con posterioridad a los eventos registrados en un vector de marcas de tiempo.

No todas las aplicaciones utilizan un registro de mensajes, algunas de ellas, incluyendo *Tattler*, pueden realizar estas operaciones directamente desde la copia del estado del grupo. En este caso la aplicación debe de proporcionar un interface o mecanismo para recuperar los mensajes junto con el identificador del host y la marca de tiempo del estado del grupo.

Vectores de marcas de tiempo de resumen

El sistema mantiene un vector de marcas de tiempo de resumen (*summary*) para registrar las últimas actualizaciones que se han dado en el sistema. El vector contiene una marca de tiempo por cada miembro del sistema. El protocolo anti-entropía con marca de tiempo garantiza que no haya "huecos"

en la secuencia de mensajes en el registro, por lo que la marca de tiempo del último mensaje indica que se han recibido todos los mensajes con marca de tiempo anteriores a este.

El vector resumen se utiliza durante una sesión de anti-entropía para determinar que mensajes no han sido recibidos todavía por un host participante de la sesión. Dos host pueden comparar sus vectores resumen para medir hasta que punto están ambos sincronizados.

Cada host también mantiene información sobre el acuse de recibo de los mensajes, aunque la información se actualiza en las sesiones de anti-entropía usando la información del vector resumen mediante un acuse de recibo resumido.

Vector de marcas de tiempo de reconocimiento

El vector de marcas de tiempo de reconocimiento se usa para registrar qué mensajes han sido reconocidos por otras replicas o host participantes del sistema. Un participante puede determinar que todos los host participantes han observado un mensaje mirando su vector de acuse de recibo local, por ejemplo, si el host A tiene una marca de tiempo t para el host B, la replica A sabe que B ha recibido todos los mensajes de cualquier remitente con una marca de tiempo menor o igual a t .

Estructura de datos como código:

```
timestampVector summary;
timestampVector ack;

typedef list ( principalId, timestamp, message, delivered ) msgList;

class msgLog {
    msgList messages;

    // manipulate messages in the log
    add(principalId, timestamp, message)
    deliver(principalId, timestamp)
    remove(principalId, timestamp, message)

    // query the log for all messages newer than some vector
    msgList listNewer(timestampVector)
    // query the log for all messages older than some timestamp
    msgList listOlder(timestamp)
    // query for all messages sent by a principal between two timestamps
    // can use special value 'ANY' for principalId
    msgList listMsgs(principalId, timestamp first, timestamp last)
}

msgLog log;
```

FIGURE 5.3: Data structures used by the TSAE communication protocol.