Consistency of Global Checkpoints Based on Characteristics of Communication Events in Multimedia Applications

Masakazu Ono and Hiroaki Higaki

Department of Computers and Systems Engineering, Tokyo Denki University
Hatoyama-cho Ishizaka, Hiki, Saitama, 350–0394, Japan
{masa, hig}@higlab.k.dendai.ac.jp

Abstract. In order achieving fault-tolerant network systems, checkpoint-recovery has been researched and many protocols have been designed. A global checkpoint taken by the protocols have to be consistent. For conventionalnetworks, a global checkpoint is defined to be consistent if there is no inconsistent message in any communication channel [1]. For multimedia communication networks, there are additional requirements for time-constrained failure-free execution and large-size message transmissions where lost of a part of the message is acceptable. In addition, based on not characteristics of data of a message but characteristics of communication events for the message, restrictions of a consistent global checkpoint are determined. This reflects deterministic and non-deterministic property of the usage of communication buffers in layered system model. This paper proposes novel criteria, for global checkpoints in multimedia communication networks.

1 Multimedia Networks

In a multimedia network, it takes longer time to transmit and receive a message. Here, the following four *pseudo events* are defined for a multimedia message $\overline{m}$ transmitted through a communication channel $\langle p_i, p_j \rangle$; $sb(\overline{m})$, $se(\overline{m})$, $rb(\overline{m})$ and $re(\overline{m})$ for start and end of sending $\overline{m}$ in p_i and start and end of receiving $\overline{m}$ in p_j, respectively. In addition, $\overline{m}$ is decomposed into a sequence $\langle pa_1, \ldots , pa_l \rangle$ of multiple *packets* for transmission. Here, $s(pa_k)$ is a *packet sending event* and $r(pa_k)$ is a *packet receipt event* for a packet pa_k. A communication event in a multimedia network is characterized by when data of a transmitted packet is determined and when data of a received packet is accepted by an application.

Packet sending events are classified into *bulky* and *stream* ones. Data of a packet sent at a bulky packet sending event is determined at beginning of sending event of a multimedia message. Here, the following properties are held:

– If $s(pa)$ for $\overline{m}$ is bulky, $e \rightarrow s(pa)$ iff $e \Rightarrow sb(\overline{m})$ where $\rightarrow$ and $\Rightarrow$ represent causal precedence and temporal precedence between two events, respectively.
– If $s(pa)$ for $\exists pa \in \overline{m}$ is bulky, $s(pa')$ for $\forall pa' \in \overline{m}$ is bulky.

M. Bubak et al. (Eds.): ICCS 2004, LNCS 3036, pp. 691–694, 2004.

On the other hand, data of a packet pa sent at a stream packet sending event is determined just at this event. Here, the following properties are held:

- If $s(pa)$ for $\overline{m}$ is stream, $e \to s(pa)$ iff $e \Rightarrow s(pa)$.
- If $s(pa)$ for $\exists pa \in \overline{m}$ is stream, $s(pa')$ for $\forall pa' \in \overline{m}$ is stream.

Same as a packet sending event, packet receipt events, are also classified into *bulky* and *stream* ones. Data of a packet received at a bulky packet receipt event is accepted at ending of receipt event of a multimedia message. Here, the following properties are held:

- If $r(pa)$ for $\overline{m}$ is bulky, $r(pa) \to e$ iff $re(\overline{m}) \Rightarrow e$.
- If $r(pa)$ for $\exists pa \in \overline{m}$ is bulky, $r(pa')$ for $\forall pa' \in \overline{m}$ is bulky.

On the other hand, data of a packet received at a stream packet sending event is accepted just at this event. Here, the following properties are held:

- If $r(pa)$ for $\overline{m}$ is stream, $r(pa) \to e$ iff $r(pa) \Rightarrow e$.
- If $r(pa)$ for $\exists pa \in \overline{m}$ is stream, $r(pa')$ for $\forall pa' \in \overline{m}$ is stream.

For a message m with conventional data, $s(m)$ and $r(m)$ are atomic. Here, each c_i in p_i is taken only when no other event occurs in p_i. However, a multimedia message is so larger that it takes longer time to transmit and receive the message. Thus, if p_i is required to take c_i during a communication event, it has to wait until an end of the event. Hence, timeliness requirement in a checkpoint protocol is not satisfied and communication overhead in recovery is increased. Therefore, c_i should be taken immediately when p_i is required to take it even during a communication event. That is, p_i sending $\overline{m} = \langle pa_1, \ldots, pa_l \rangle$ takes c_i between $s(pa_s)$ and $s(pa_{s+1})$ and p_j receiving $\overline{m}$ takes c_j between $r(pa_r)$ and $r(pa_{r+1})$. In addition, part of a multimedia message may be lost in a communication channel for an application. Such an application requires not to retransmit lost packets in recovery but to transmit packets with shorter transmission delay. Hence, an overhead for taking a checkpoint during failure-free execution is required to be reduced.

2 Consistency in Multimedia Networks

Global consistency Gc denotes degree of consistency for a global checkpoint $C_{\mathcal{V}} = \{c_1, \ldots, c_n\}$. In a conventional network, Gc is defined as follows:

$$Gc = \begin{cases} 1 & \text{no inconsistent message.} \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

In a multimedia network, a local checkpoint is taken even during a communication event and it is acceptable to lose part of a multimedia message. Hence, a domain of Gc is a closed interval $[0, 1]$ instead of a discrete set $\{0, 1\}$.

2.1 Message Consistency

Message consistency. Mc^u_{ij} is degree of consistency for a set $\{c_i, c_j\}$ of local checkpoints and a multimedia message $\overline{m_u}$ through a communication channel $\langle p_i, p_j \rangle$. Here, we define an *inconsistent multimedia message*.

[Inconsistent multimedia message.] m_u is inconsistent iff m_u is a *lost* or an *orphan multimedia message*. m_u is a lost multimedia message iff $se(\overline{m_u})$ occurs before c_i in p_i and $rb(\overline{m_u})$ occurs after c_j in p_j. m_u is an orphan multimedia message iff $sb(\overline{m_u})$ occurs after c_i in p_i and $rb(\overline{m_u})$ occurs before c_j in p_j. □
[Consistency for inconsistent multimedia message.] $Mc^u_{ij} = 0$ for an inconsistent multimedia message $\overline{m_u}$. □

Next, inconsistent packets are introduced due to checkpoints during a communication event.

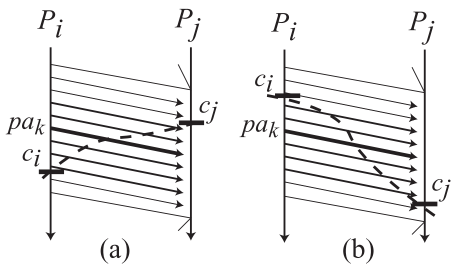

Fig. 1. Lost packet (a) and Orphan packet (b).

[Lost and orphan packets.] Suppose that c_i and c_j are between $s(pa_s)$ and $s(pa_{s+1})$ and between $r(pa_r)$ and $r(pa_{r+1})$ for $\overline{m_u} = \langle pa_1, \ldots, pa_l \rangle$, respectively. pa_k is a *lost packet* iff $s(pa_k)$ occurs before c_i in p_i and $r(pa_k)$ occurs after c_j in p_j. pa_k is an *orphan packet* iff $s(pa_k)$ occurs after c_i in p_i and $r(pa_k)$ occurs before c_j in p_j.□

If $s = r$, there is no lost and orphan packet. Hence, $Mc^u_{ij} = 1$.

If $s > r$, $\{pa_{r+1}, \ldots, pa_s\}$ is a set of lost packets. These packets are not retransmitted after recovery. Lost packets in a conventional network are restored by logging them in failure-free execution. However, in a multimedia network, less overhead in failure-free execution is required since applications require time-constrained execution. In addition, even if part of a multimedia message is lost in recovery, an application accepts the message. The less packets are lost, the higher message consistency is achieved. A multimedia message is usually compressed for transmission. Thus, value of packets for a message is not unique. Therefore, message consistency depends on total value of lost packets as follows:

$$\frac{\partial Mc^u_{ij}}{\partial l\text{-}value} < 0 \quad \text{where } l\text{-}value = \sum_{lost\ packets\ pa_k} value(pa_k). \tag{2}$$

Here, a domain of Mc^u_{ij} is an open interval $(0, 1)$.

If $s < r$, $\{pa_{s+1}, \ldots, pa_r\}$ is a set of orphan packets. An orphan multimedia message might not be retransmitted after recovery. However, orphan packets are surely retransmitted since c_i and c_j are taken during transmission and receipt of $\overline{m_u}$ and the data of $\overline{m_u}$ being carried by a sequence $\langle pa_1, \ldots, pa_l \rangle$ of packets is not changed even after recovery. Hence, message consistency does not depend on orphan packets.

[**Message consistency**] Let $value(\overline{m_u})$ be total value of all packets in $\overline{m_u}$.

$$\begin{aligned} Mc_{ij}^u &= 0 && \text{if } l\text{-}value = value(\overline{m_u}). \\ Mc_{ij}^u &= 1 && \text{if } l\text{-}value = 0. \\ \frac{\partial Mc_{ij}^u}{\partial l\text{-}value} &< 0 && \text{otherwise.} \quad \square \end{aligned} \tag{3}$$

2.2 Channel Consistency

Channel Consistency. Cc_{ij} is calculated by Mc_{ij}^u for every message $\overline{m_u}$ through $\langle p_i, p_j\rangle$. For compatibility with (1), if message consistency for every message through $\langle p_i, p_j\rangle$ is 1, channel consistency is also 1. On the other hand, if message consistency for at least one message through $\langle p_i, p_j\rangle$ is 0, channel consistency is also 0. In addition, channel consistency monotonically increases for consistency of messages through $\langle p_i, p_j\rangle$.

[**Channel consistency.**] Let $\mathcal{M}_{ij}$ be a set of messages through $\langle p_i, p_j\rangle$.

$$\begin{aligned} Cc_{ij} &= 1 && \text{if } \forall \overline{m_u} \in \mathcal{M}_{ij} \quad Mc_{ij}^u = 1. \\ Cc_{ij} &= 0 && \text{if } \exists \overline{m_u} \in \mathcal{M}_{ij} \quad Mc_{ij}^u = 0. \\ \forall \overline{m_u} \in \mathcal{M}_{ij},\ \frac{\partial Cc_{ij}}{\partial Mc_{ij}^u} &> 0 && \text{otherwise.} \quad \square \end{aligned} \tag{4}$$

2.3 Global Consistency

Gc is calculated by Cc_{ij} for every communication channel $\langle p_i, p_j\rangle$. For compatibility with (1), if channel consistency for every channel is 1, global consistency is also 1. On the other hand, if consistency for at least one communication channel is 0, global consistency is also 0. In addition, channel consistency monotonically increases for consistency of the communication channels.

[**Global consistency.**]

$$\begin{aligned} Gc &= 1 && \text{if } \forall\langle p_i, p_j\rangle \quad Cc_{ij} = 1. \\ Gc &= 0 && \text{if } \exists\langle p_i, p_j\rangle \quad Cc_{ij} = 0. \\ \forall\langle p_i, p_j\rangle,\quad \frac{\partial Gc}{\partial Cc_{ij}} &> 0 && \text{otherwise.} \quad \square \end{aligned} \tag{5}$$

3 Conclusion

This paper has proposed a novel criteria for consistency of a global checkpoint in multimedia network systems. The authors will design QoS-based checkpoint protocols based on the criteria.

References

1. Chandy, K.M. and Lamport, L., "Distributed Snapshots: Determining Global States of Distributed Systems," ACM Trans. on Computer Systems, Vol. 3, No. 1, pp. 63–75 (1985).