

## チェックポイントプロトコル実行中のトポロジ変化を考慮した無線マルチホップネットワークのためのチェックポイントプロトコル

東京電機大学 理工学部 情報システム工学科

小野 真和 桧垣 博章 足立 暁生

E-mail: {masa,hig}@higlab.k.dendai.ac.jp, adachi@k.dendai.ac.jp

分散コンピューティング環境において耐故障性を実現する手法として、従来の固定ネットワーク環境ではチェックポイントプロトコルが提案されてきた。チェックポイントプロトコルでは、状態情報を格納する安定記憶の存在と、メッセージの送信元コンピュータと送信先のコンピュータの同期による一貫性のないメッセージ(紛失メッセージ、孤児メッセージ)の検出、回避が十分に可能な通信帯域の存在が前提となっている。また、メッセージの配送は送信元のアプリケーションから送信先のアプリケーションまで直接配送可能であるとしている。本論文では、これらの前提条件が成立しない、移動コンピュータのみで構成され、メッセージがマルチホップで配送されるようなアドホックネットワークに注目し、アドホックネットワークにおけるチェックポイントリカバリプロトコルを提案する。本提案では移動コンピュータの状態情報を複数の隣接移動コンピュータに保存する。また、転送中に送信先において紛失メッセージとなる可能性のあるメッセージを紛失可能性メッセージとして中継移動コンピュータで保存する。このとき、中継移動コンピュータの状態情報の一部として記憶することにより、状態情報とメッセージログを同一の移動コンピュータに同時に保存することができる。これによって、チェックポイントプロトコルの開始から終了までに要する時間を短縮することが可能である。

## Checkpoint Protocol for Mobile Ad Hoc Network

Masakazu Ono, Hiroaki Higaki and Akeo Adachi

Department of Computers and Systems Engineering

Tokyo Denki University

E-mail: {masa,hig}@higlab.k.dendai.ac.jp, adachi@k.dendai.ac.jp

For achieving mission-critical network applications, checkpoint recovery protocols have been researched and developed. In conventional protocols for wired networks, stable storages to store state information are assumed and enough bandwidth is assigned to synchronize a sender and a receiver computers of a message in order to avoid that the message becomes inconsistent, i.e. neither orphan nor lost. In this paper, we propose a novel checkpoint protocol in ad hoc networks without stable storage and enough communication bandwidth. Here, a checkpoint request message is delivered by flooding. State information of a mobile computer is carried by this message and stored into neighbor mobile computers. A candidate of a lost message is detected and stored by intermediate mobile computer on its transmission route. Here, communication overhead for taking global checkpoint is reduced.

## 1 はじめに

ノート型コンピュータやPDAなどの移動コンピュータをIEEE802.11 [3]やHIPERLAN [1]、Bluetooth [2]などの無線通信プロトコルを用いて相互に接続した無線LAN技術の研究開発が進み、その使用が普及している。また、無線基地局を介して有線ネットワークと接続されたインフラストラクチャネットワークだけでなく、移動コンピュータだけで構成されるアドホックネットワークの研究も進んでいる。アドホックネットワークの適用例として、一時的に構築されるイベント会場や災害現場などでのネットワーク、危険地帯で無線基地局が設置できない場所における自律移動型ロボットの協調動作のためのネットワーク、センサネットワーク [5,14]などがある。このようなアドホックネットワークにおけるミッションクリティカルアプリケーションの実行を考えたとき、耐故障性を実現するためのチェックポイントリカバリ手法を適用することが考えられる。しかし、有線ネットワーク環境を対象とした従来のチェックポイント手法 [9]は安定記憶 [13]がネットワーク上に存在することを前提としている。また、一貫性のないメッセージ(孤児メッセージと紛失メッセージ)を送信元コンピュータと送信先コンピュータの同期によって検出することが可能となる十分な帯域幅がネットワークによって提供されているとしている。そのため、アドホックネットワーク上の移動コンピュータは安定記憶を持つことができない問題や、アドホックネットワークにおける隣接移動コンピュータ間の通信リンクが狭帯域幅で低信頼であるため、このネットワークを介した同期に要する通信オーバーヘッドが大きい問題を解決する必要がある。また、メッセージの配送が送信元のアプリケーションから送信先のアプリケーションまで複数の移動コンピュータを経由して配送される、マルチホップ配送を行なっている。このとき、移動コンピュータの移動などによってメッセージの配送経路が変わり、従来とは異なった条件で一貫性のないメッセージが発生する。そこで、本論文ではアドホックネットワークにおいて安定記憶を実現し、一貫性のないメッセージの発生を回避するための送受信コンピュータ間における同期の実現に起因する通信オーバーヘッドを回避する新たなチェックポイントプロトコルを提案する。

## 2 従来手法

### 2.1 チェックポイントプロトコル

アドホックネットワーク  $\mathcal{N} = (\mathcal{V}, \mathcal{E})$  とは、移動コンピュータの集合  $\mathcal{V}$  と互いに直接メッセージを交換することが可能な移動コンピュータ  $M_i, M_j \in \mathcal{V}$  の間の双方向リンク  $\langle M_i, M_j \rangle$  の集合  $\mathcal{E}$  で定まるネットワークである。一般に、ネットワーク環境において、チェックポイントプロトコルによって各移動コンピュータ  $M_i \in \mathcal{V}$  が設定したローカルチェックポイント  $c_i$  の集合である

グローバルチェックポイント  $C_V$  が一貫性を持つとは、次の性質を満たすことをいう [6]。

[定義]

- 1) 送信元移動コンピュータ  $M_s$  から送信先移動コンピュータ  $M_r$  へ配送されるメッセージ  $m$  が紛失メッセージであるとは、グローバルチェックポイント  $C_V$  に対して、 $Send(m)$  が  $c_s$  に先行し、 $c_r$  が  $Receive(m)$  に先行することである。なお、 $Send()$  と  $Receive()$  は、アプリケーション層におけるメッセージ送受信イベントである。
- 2) メッセージ  $m$  が孤児メッセージであるとは、 $C_V$  に対して、 $c_s$  が  $Send(m)$  に先行し、 $Receive(m)$  が  $c_r$  に先行することである。
- 3) 紛失メッセージ、孤児メッセージを含まないグローバルチェックポイントは、一貫性が保たれているという。□

ただし、紛失メッセージをリカバリ回復時に再送信することができれば、システム状態の一貫性を維持することが可能である。そこで、一貫性のあるグローバルチェックポイントを以下のように再定義する。

[定義]

- 4) 一貫性のあるグローバルチェックポイントとは、孤児メッセージを含まず、すべての紛失メッセージをリカバリ回復時に再送信可能であるものである。□

この定義にしたがって、 $M_r$  で紛失メッセージ  $m$  を記憶、保存するプロトコルとして [8] がある。

従来のチェックポイントプロトコルにおいては、 $m$  が  $C_V$  に対する紛失メッセージや孤児メッセージとなることを  $M_r$  でのみ判定することを前提としている。そのため、これらの発生を回避するには、システム全体での同期を必要としていた。例えば、Koo [12] のプロトコルにおいては、チェックポイント要求メッセージ  $CReq$  を受信してから、チェックポイント終了メッセージ  $CFin$  を受信するまでの間、アプリケーションメッセージの送信を禁止している。ところが、アドホックネットワークにおいては、無線通信の狭帯域幅、無線信号の減衰と複数無線信号の衝突による低信頼性、無線通信帯域の予約における競合、マルチホップ配送による伝達遅延などのために移動コンピュータ間の同期に要する通信オーバーヘッドが大きくなる。この結果、アプリケーションの実行を一時停止する時間が長くなる、すなわち、チェックポイントプロトコルの開始から終了までの時間が長くなるという問題が発生する。本論文で提案するプロトコルでは、アドホックネットワークにおいて、 $m$  が紛失メッセージあるいは孤児メッセージとなる可能性を  $m$  の配送経路上にある移動コンピュータが判定し、必要に応じて  $m$  を記憶したり、 $m$  の転送を遅延させたりすることによってこの問題を解決する。ここでは、隣接する移動

コンピュータ間における同期のみが必要であることから、アプリケーションの停止時間を短縮することが可能である。

## 2.2 モバイルチェックポイントプロトコル

モバイルチェックポイントプロトコルの実現にあたり、論文 [15] では、移動コンピュータを含むネットワークを以下の4つのモデルに分類している。

- 1) Centralized Networks
- 2) Cell-Dependent Infrastructured Networks
- 3) Cell-Independent Infrastructured Networks
- 4) Ad-hoc Networks

1)-3) は、ネットワークの構成要素に固定コンピュータを含んでいる。そこで、固定コンピュータに実現した安定記憶に移動コンピュータの状態情報を保存することにより、チェックポイントを設定することができる。論文 [10] では、同期チェックポイント手法と非同期チェックポイント手法を組み合わせた複合チェックポイント手法を提案している。[10], [16] および [15] において、それぞれ 1), 2) に対するプロトコルを設計している。また、[4] と [17] も 1) を対象として固定コンピュータの安定記憶を使用するプロトコルを提案している。ところが 4) においては、ネットワークの構成要素が移動コンピュータのみであることから、安定記憶の実現が困難である。そこで、各移動コンピュータのチェックポイントの設定を、複数の移動コンピュータに状態情報を記憶させることによって実現する。

## 3 提案プロトコル

### 3.1 チェックポイントプロトコル

以下の条件のもとでプロトコルを構成する。

[前提条件]

- 1) アドホックネットワークに含まれるすべての移動コンピュータは、チェックポイントプロトコルの実行中、マルチホップ配送により互いにメッセージ交換が可能である。
- 2) 隣接する移動コンピュータ間の通信リンクは、動的に切断および確立されることがある。
- 3) 各移動コンピュータは、隣接する移動コンピュータのリストを保持している。
- 4) 隣接する移動コンピュータ間の通信リンクは双方向であり、半二重通信が行なわれる。また、ユニキャスト通信は、受信確認と再送機構により、メッセージの紛失なく実現されているものとする。□

チェックポイントプロトコルの基本形を示す。チェックポイントプロトコルの開始は、任意の移動コンピュータが任意のタイミングで行なうことができる。チェックポイント設定要求の伝達と、チェックポイント間の同期は、チェックポイント設定要求メッセージ  $CReq$  のフラッ

ディング [7] によって実現される (図 1)。

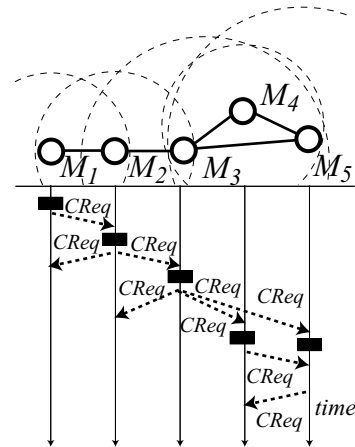


図 1: チェックポイントプロトコル

$CReq$  を受信した移動コンピュータ  $M_i$  は、現在の状態情報  $S_i$  を獲得することによってローカルチェックポイント  $c_i$  を設定するとともに、 $CReq$  を隣接する移動コンピュータ群、すなわち、 $M_i$  の無線信号到達範囲内に存在するすべての移動コンピュータにブロードキャストする。これを繰り返すことによって、前提条件 1) により、すべての移動コンピュータにおいて、ローカルチェックポイントを設定することができる。

ここで、移動コンピュータ  $M_i$  の状態情報  $S_i$  を保存する安定記憶を実現するために、 $S_i$  を複数の隣接移動コンピュータに記憶させる手法を用いる。各移動コンピュータは、ローカルチェックポイント  $c_i$  における状態情報  $S_i$  を獲得してから  $CReq$  のブロードキャストを行なうことから、 $S_i$  を  $CReq$  によって配送することにより、追加のメッセージを要することなく  $S_i$  の配送が実現される。

[アドホックチェックポイントプロトコル (基本形)]

- 1) 任意の移動コンピュータ  $M_0$  が、 $M_0$  の状態情報  $S_0$  を獲得することでローカルチェックポイント  $c_0$  を設定するとともに、 $S_0$  を含み、 $M_0$  が生成した ID が付与されたチェックポイント設定要求メッセージ  $CReq$  を  $M_0$  の無線信号到達範囲内にブロードキャストする。このとき、タイマ  $T_0$  をセットする。
- 2) 移動コンピュータ  $M_i$  が送信したチェックポイント設定要求メッセージ  $CReq$  を受信した移動コンピュータ  $M_j$  は、以下の処理を行なう。
  - 2-1)  $M_i$  から同一の ID を持つ  $CReq$  を受信していないならば、受信した  $CReq$  に含まれる  $M_i$  の状態情報  $S_i$  を保存する。
  - 2-2)  $M_j$  がいずれの隣接移動コンピュータからも同一の ID を持つ  $CReq$  を受信していないならば、 $M_j$  の状態情報  $S_j$  を獲得するとともに、 $S_j$  を含み、受信した  $CReq$  と同一の ID を付与した

$CReq$  を  $M_j$  の無線信号到達範囲内にブロードキャストする。このとき、タイマ  $T_j$  をセットする。

- 3) 移動コンピュータ  $M_j$  が隣接移動コンピュータリスト  $L_j$  に含まれるすべての移動コンピュータから  $CReq$  を受信する以前にタイマ  $T_j$  が時間切れとなったならば、 $M_j$  は、同じ  $CReq$  を再度ブロードキャストする。□

ここで、チェックポイントプロトコルの実行と並行に送受信されたメッセージは、紛失メッセージや孤児メッセージとなる可能性がある。紛失メッセージは、いずれかの移動コンピュータに保存し、リカバリ回復時に、保存されたメッセージを再送信することによって、システム状態の一貫性を維持することができる。一方、孤児メッセージは、リカバリ再実行時に送信元移動コンピュータが同一のメッセージを再度送信する保障がないことから、その発生を回避しなければならない。移動コンピュータの移動による通信路の切断と接続が発生しない場合には、以下の性質が成り立つ。

[性質]

- 1) 紛失メッセージ  $m_i$  は、その配送経路上で以下のいずれかの条件を満足する。
  - 1-a)  $m_i$  の配送経路上にある 1 台以上の移動コンピュータ  $M_i$  において、 $receive(m_i) \rightarrow c_i \rightarrow send(m_i)$  が成り立つ。  
ただし、 $send()$  と  $receive()$  は、ネットワーク層における送受信イベントである。また、 $\rightarrow$  は、イベント間の happened-before の関係を表すものとする。
  - 1-b)  $m_i$  の配送経路上にある 2 台の移動コンピュータ  $M_i, M_j$  において、 $M_i, M_j$  は互いに直接通信可能であり、 $send(m_i) \rightarrow c_i$  かつ  $c_j \rightarrow receive(m_i)$  が成り立つ。
- 2) 孤児メッセージ  $m_o$  は、その配送経路上で以下の条件を満足する。
  - 2-a)  $m_o$  の配送経路上にある 2 台の移動コンピュータ  $M_i, M_j$  において、 $M_i$  と  $M_j$  は互いに直接通信可能であり、 $c_i \rightarrow send(m_o)$  かつ  $receive(m_o) \rightarrow c_j$  が成り立つ。□

性質 1) により、紛失メッセージとなる可能性があるメッセージ  $m_i$  を中継移動コンピュータ、すなわち  $m_i$  の送信元でも送信先でもない移動コンピュータを検出することができる。もし、この検出が送信先移動コンピュータ  $M_r$  でのみ検出可能であるならば (例えば、論文 [6], [12] では、 $M_r$  で  $m_i$  が紛失メッセージとなることを検出している。)、図 2 に示すように、 $M_r$  が  $m_i$  を受信した時点、すなわち  $Receive(m_i)$  においては、 $M_r$  がすでに  $CReq$  を隣接移動コンピュータへブロードキャスト送信済みであることが考えられる。この場合、 $m_i$

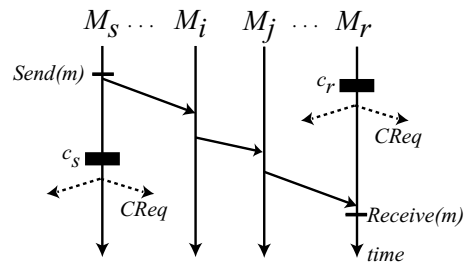


図 2: 紛失メッセージの検出とチェックポイント

を隣接移動コンピュータに記憶させるために、 $m_i$  を含む制御メッセージをブロードキャストする必要がある。しかし、これによって、以下の問題が発生する。

- (1)  $M_r$  がリカバリで必要とする情報を隣接移動コンピュータに記憶させるために  $M_r$  が送受信するメッセージが増加する。
- (2)  $M_r$  の状態情報  $S_r$  を保存した隣接移動コンピュータがこの制御メッセージの送信時点においても  $M_r$  の無線信号到達範囲内に存在することは保証できない。すなわち、 $S_r$  のみを持つ移動コンピュータ、 $m_i$  のみを持つ移動コンピュータが発生することがある。この結果、リカバリで必要とする情報が複数の移動コンピュータに分散することによって、リカバリ時に送受信されるメッセージ数が増加する。
- (3) リカバリ回復時の再送信を必要とするすべての紛失メッセージを隣接移動コンピュータに記憶し、チェックポイントプロトコルが終了したことを各移動コンピュータが検出するためには、新しい同期メッセージの導入が必要となる。

そこで、紛失メッセージとなる可能性のあるメッセージ  $m_i$  を  $CReq$  を送信する前に検出可能な移動コンピュータの存在が不可欠である。ここで、条件 1-a) を満たすメッセージ  $m_i$  については、移動コンピュータ  $M_i$  が  $m_i$  を  $CReq$  の送信前に検出することができる。

[紛失メッセージとなる可能性の検出 (1)]

移動コンピュータ  $M_i$  が中継メッセージ ( $M_i$  を送信元、送信先としないメッセージ)  $m_i$  を  $CReq$  の送信前に受信し、 $CReq$  送信時までに送信していないならば  $m_i$  は紛失メッセージとなる可能性のあるメッセージである。 $M_i$  は状態情報  $S_i$  とともに  $m_i$  を  $CReq$  に含めてブロードキャストし、これを受信した  $M_i$  の隣接移動コンピュータは、 $S_i$  と  $m_i$  を記憶する。□

一方、条件 1-b) を満たすメッセージ  $m_i$  については、 $m_i$  が紛失メッセージとなる可能性があるメッセージであることを検出できるのは  $M_j$  であり、検出したとき  $M_j$  はすでに  $CReq$  メッセージを送信済みであることがある。そこで、以下の方法により  $m_i$  の検出を  $M_j$  が行ない、 $M_i$  が  $CReq$  を送信する前にその結果を  $M_i$  に通

知る (図 3)。

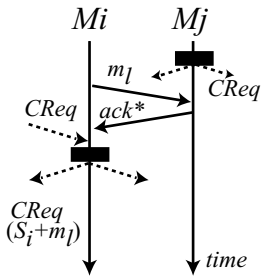


図 3: 中継移動コンピュータでの紛失メッセージの検出とチェックポイント

[紛失メッセージとなる可能性の検出 (2)]

$CReq$  を送信していない移動コンピュータ  $M_i$  から送信されたメッセージ  $m_i$  を移動コンピュータ  $M_j$  が  $CReq$  送信後に受信したならば、 $m_i$  の受信確認応答メッセージに  $m_i$  が紛失メッセージとなる可能性があることを示す情報を付加する。 $M_i$  は、これを受信したならば、状態情報  $S_i$  とともに  $m_i$  を  $CReq$  に含めてブロードキャストする。この手法を適用するために、各移動コンピュータは、メッセージ送信時から受信確認応答メッセージの受信までの間は  $CReq$  を送信することができない。□

条件 1-a)、条件 1-b) はメッセージ  $m_i$  が紛失メッセージとなる必要条件であり、十分条件ではない。例えば、図 4 において、 $M_1$  では  $send(m) \rightarrow c_1$  であり、 $M_2$  では  $c_2 \rightarrow receive(m)$  であることから、 $M_2$  によって紛失メッセージとなる可能性があると判定される。この結果、 $m$  は  $M_1$  からブロードキャストされる  $CReq$  に含まれる。しかし、 $M_1$  では  $Send(m) \rightarrow c_1$  であり、 $M_3$  では  $Receive(m) \rightarrow c_3$  であることから、 $m$  は紛失メッセージではない。

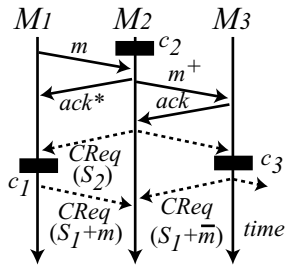


図 4: 誤検出紛失メッセージの多重受信回避 (誤検出の場合)

もし、グローバルチェックポイント  $C_V = \{c_1, c_2, c_3\}$  からリカバリしたならば、 $m$  は  $M_1$  によって再送信されるが、 $M_3$  では受信済みである。ここで、 $m$  が紛失メッセージとなる可能性のあるメッセージであることを検出した移動コンピュータである  $M_2$  ( $m$  を含む  $CReq$  を送信した移動コンピュータである  $M_1$  ではない) は、 $m$  が

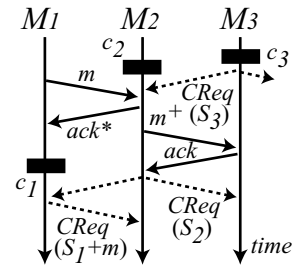


図 5: 誤検出紛失メッセージの多重受信回避 (誤検出でない場合)

リカバリ回復時に再送信されるメッセージであることを示す情報を  $m$  に付与する。 $m$  を受信した送信先移動コンピュータ  $M_3$  では、 $m$  の受信が  $CReq$  の送信前であるならば、再送信時に  $m$  を受信しても破棄することを示す情報を  $CReq$  に含めることができる (図 4)。また、 $m$  の受信が  $CReq$  の送信後であるならば、 $m$  は紛失メッセージであるので、リカバリ回復後に再送信されたものを受信する必要がある (図 5)。

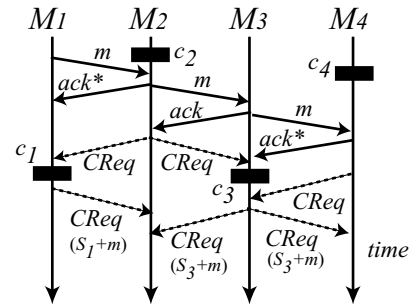


図 6: 紛失メッセージの多重検出

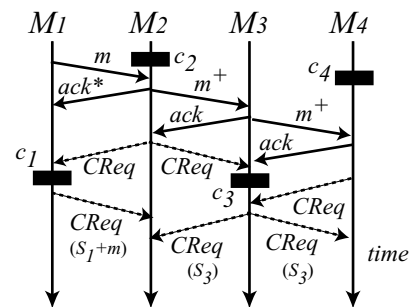


図 7: 紛失メッセージの多重検出の回避

また図 6 のように複数の移動コンピュータが、1 つのメッセージを紛失メッセージとなる可能性のあるメッセージであると検出することがある。例えば、 $m$  は  $M_2$  および  $M_4$  において紛失メッセージとなる可能性のあるメッセージとして検出される。

これは、 $M_1$  において  $send(m) \rightarrow c_1$  であり、 $M_2$  において  $c_2 \rightarrow receive(m)$  であることと、 $M_3$  において

$send(m) \rightarrow c_3$  であり、 $M_4$  において  $c_4 \rightarrow receive(m)$  であることによるものである。もし、これらの移動コンピュータがそれぞれの  $CReq$  に  $m$  を含めるならば、リカバリ時に送信先移動コンピュータ  $M_4$  は  $M_1$  と  $M_3$  から再送信された 2 つの  $m$  を受信することになる。この問題を解決するために、前段で述べた再送信時に破棄すべきメッセージであることを示す情報を用いる (図 6)。もし、この情報を含むメッセージを紛失メッセージの可能性のあるメッセージであると検出しても、これを以降に送信される  $CReq$  に含めないことによって、最初に紛失メッセージとなる可能性を検出した移動コンピュータ (条件 1-a) を満たす場合) の送信する  $CReq$ 、または、この移動コンピュータにこのメッセージを送信した移動コンピュータ (条件 1-b) を満たす場合) の送信する  $CReq$  にのみ  $m$  を含ませることができる。

孤児メッセージとなる可能性のあるメッセージは、中継移動コンピュータでは対処を行わない。送信先移動コンピュータで受信を遅延させることによってのみ対処すれば十分である。

次に、チェックポイントプロトコル実行中のトポロジ変化について考える。アドホックネットワークでは移動コンピュータ間の通信リンクが低信頼であり、ノイズなどの外乱によって通信リンクが切断されたり、移動コンピュータの移動によって通信リンクが変わるといった特徴がある。ただし、本論文ではネットワークの分割や合流も考慮するが、ネットワークが分割された状態となる時間は短く、分割している間に複数回チェックポイントが行なわれたり、分割したネットワークごとに異なるチェックポイントが実行される環境ではないとする。

このように、トポロジが変化するような場合、以下の問題が考えられる。

[トポロジ変化によって発生する問題]

- 1) ローカルチェックポイントの設定が行なえない移動コンピュータが存在する
- 2) リンク接続時に、送信したメッセージが保存されなのまま紛失メッセージとなる場合がある

問題 1) について、プロトコルの開始前または実行中にネットワークが分割したとき、ネットワーク内の各コンピュータがチェックポイントの設定を完了しているネットワークと、チェックポイントの設定を行っていないネットワークが構成される場合がある。このようなネットワークが合流し、1 つのネットワークになったとしても、前項で提案したプロトコルではチェックポイントの設定が設定前後のネットワークの合流が発生した際の  $CReq$  の再送についてプロトコルでは定義されていない。

そこで、チェックポイント設定済みの移動コンピュータ  $M_i$  とチェックポイントの設定を行っていない移動コン

ピュータ  $M_j$  間でリンクを確立させる場合には、その前に  $M_j$  がチェックポイントの設定を開始するものとする。たとえば、図 8 は、あるネットワークが  $M_1, M_2, M_3, M_4$  から構成されるネットワークと  $M_5, M_6, M_7$  から構成されるネットワークに分割されている。また、 $M_5, M_6, M_7$  のネットワークではチェックポイントの設定がすべて終了しているとする。時刻  $t$  において、 $M_4$  と  $M_5$  の間でリンクが確立したとき、 $M_5$  はチェックポイントが完了し、 $M_4$  ではチェックポイントの設定が行なわれる前の状態である。そこで、 $M_4$  はチェックポイントプロトコルを開始する。 $M_4$  がチェックポイントプロトコルを開始することによって、 $M_1, M_2, M_3$  もチェックポイントを設定することができる。以上のことから、この手法を利用することによりトポロジが変化する場合でもすべての移動コンピュータがチェックポイントを設定することができる。

問題 2) について、チェックポイントの設定が行なわれていない移動コンピュータ  $M_j$  からチェックポイントの設定が完了している移動コンピュータ  $M_i$  にメッセージ  $m$  が送信される場合送信先において紛失メッセージとなる場合がある。また、送信先の移動コンピュータまでの経路でチェックポイントプロトコルが実行される前のコンピュータの存在を保証することができない。そこで、 $M_j$  がリンクの確立前に実行するチェックポイントの設定時に、 $M_j$  が転送する予定であるメッセージのうち紛失となる可能性のあるメッセージを  $S_j$  に付加することとする。図 8 において、メッセージ  $m$  が  $M_1$  から  $M_6$  に送信される場合を考える。また、時刻  $t$  よりも以前に  $m$  は  $M_4$  に転送されているとする。リンク確立後、 $m$  が  $M_6$  に転送されたとき、 $m$  は紛失メッセージであると検出される。しかし、 $M_5, M_6, M_7$  においてはすでにチェックポイントプロトコルが終了してしまっている。そこで、 $M_4$  はリンク確立時、チェックポイントプロトコルを実行するときに  $m$  を状態情報に付加することができ、そのようにすることによってメッセージの再送信を保証することができる。したがって、紛失の可能性のあるメッセージを保存し、リカバリ時にも再送信できることからグローバルチェックポイントの一貫性を維持することができる。

以上の点から、トポロジが変化する場合においても、すべてのコンピュータがチェックポイントプロトコルを実行することができ、また紛失となる可能性のあるメッセージを保存することができることから一貫性のあるグローバルチェックポイントを定めることができる。

### 3.2 プロトコル

以下にメッセージ通信プロトコルおよびチェックポイントプロトコルを示す。

[通信リンク確立時]

移動コンピュータ  $M_i$  と  $M_j$  の通信リンク確立時、互いにチェックポイント設定済みの  $ID$  を交換する。もし、 $M_i$

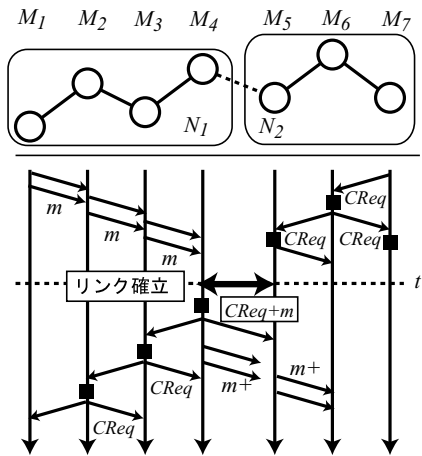


図 8: トポロジ変化によって発生するチェックポイントの実行とメッセージの保存

および  $M_j$  とともにチェックポイントプロトコルを実行していないのならば  $ID_i < ID_j$  であれば、 $M_i$  がチェックポイントプロトコルを開始し、 $ID_i > ID_j$  であれば、 $M_j$  がチェックポイントプロトコルを開始する。

このとき、自身が持つ転送しなければならないメッセージのうち、紛失となる可能性のあるメッセージを状態情報に付加する。

[メッセージ通信プロトコル]

移動コンピュータ  $M_i$  は、アプリケーションプログラムにおける送信要求、受信要求が発生したならば、アプリケーション層の送信イベント  $Send()$ 、受信イベント  $Receive()$  を実行する。また、ネットワークからメッセージ  $m$  を受信したならば  $receive()$  を実行する。 $send()$  は、送信元移動コンピュータにおける  $Send()$  実行時と中継移動コンピュータにおける  $receive()$  実行時に実行される。メッセージ  $m$  には、 $m.source\_creq\_sent$ 、 $m.creq\_sent$ 、 $m.logged$  という 3 つのフラグが含まれる。

(アプリケーション層の送信イベント  $Send(m)$ )

- 1  $m.logged := false$  とする。
- 2  $M_i$  が  $CReq$  を送信済みであるならば、 $m.source\_creq\_sent := true$ 、未送信であるならば  $m.source\_creq\_sent := false$  とする。
- 3  $send(m)$  を実行する。

(アプリケーション層の受信イベント  $Receive(m)$ )

- 1  $m$  が  $buffer_i$  に含まれないならば、 $receive(m)$  の実行が終了するまで一時停止する。
- 2 もし、 $m.source\_creq\_sent = true$  かつ  $M_d$  が  $CReq$  を未送信であるならば、 $CReq$  の送信が終了するまで一時停止する。
- 3  $m$  を  $buffer_i$  から取り出す。
- 4  $m.logged = true$  かつ、 $M_i$  が  $CReq$  を未送信であるならば、リカバリ回復後に受信する  $m$  を破棄す

る情報を以降に送信される  $CReq$  に含める。□

(ネットワーク層の送信イベント  $send(m)$ )

- 1  $M_i$  が  $CReq$  を送信済みであるならば、 $m.creq\_sent := true$ 、未送信であるならば  $m.creq\_sent := false$  とする。
- 2  $CReq$  の送信を不可とする。
- 3  $m$  を送信する。
- 4  $ack(m)$  を受信する。 $m.logged = false$  かつ  $ack(m).logged = true$  であるならば、以降に送信される  $CReq$  に  $m$  を含める。
- 5  $CReq$  の送信を可とする。

(ネットワーク層の受信イベント  $receive(m)$ )

- 1  $m.logged = false$  かつ  $m.creq\_sent = false$  かつ  $M_i$  が  $CReq$  を送信済みであるならば、 $m.logged := true$ 、 $ack(m).logged := true$  として  $ack(m)$  を返送する。
- 2 それ以外の場合は、 $ack(m).logged := m.logged$  として  $ack(m)$  を返送する。
- 3  $m$  の送信先が  $M_i$  であるならば、 $m$  を  $buffer_i$  に格納する。
- 4 それ以外の場合は  $send(m)$  を実行する。

各移動コンピュータ  $M_i$  は、 $CReq$  に状態情報  $S_i$  とともに含ませるメッセージの集合をメッセージログ  $ML_i$  として保存することとする。

[アドホックチェックポイントプロトコル]

- 1) 任意の移動コンピュータ  $M_0$  が、 $M_0$  の状態情報  $S_0$  を獲得することでローカルチェックポイント  $c_i$  を設定するとともに、 $S_0$  とメッセージログ  $ML_0$  を含み、 $M_0$  が生成した  $ID$  が付与されたチェックポイント設定要求メッセージ  $CReq$  を  $M_0$  の無線信号到達範囲内にブロードキャストする。このとき、タイマ  $T_0$  をセットする。
- 2) 移動コンピュータ  $M_i$  が送信したチェックポイント設定要求メッセージ  $CReq$  を受信した移動コンピュータ  $M_j$  は、以下の処理を行なう。
  - 2-1)  $M_i$  から同一の  $ID$  を持つ  $CReq$  を受信していないならば、受信した  $CReq$  に含まれる  $M_i$  の状態情報  $S_i$  とメッセージログ  $ML_i$  を保存する。
  - 2-2)  $M_j$  がいずれの隣接移動コンピュータからも同一の  $ID$  を持つ  $CReq$  を受信していないならば、 $M_i$  の状態情報  $S_i$  とメッセージログ  $ML_i$  を獲得するとともに、 $S_j$  を含み、受信した  $CReq$  と同一の  $ID$  を付与した  $CReq$  を  $M_j$  の無線信号到達範囲内にブロードキャストする。このとき、タイマ  $T_j$  をセットする。
- 3) 移動コンピュータ  $M_j$  が近隣する移動コンピュータリスト  $L_j$  に含まれるすべての移動コンピュータから  $CReq$  を受信する以前にタイマ  $T_j$  が時間切れとなったならば、 $M_j$  は、同じ  $CReq$  を再度ブロード

キャストする。

- 4)  $ML_i := \phi$  として終了する。□

#### 4 まとめと今後の課題

本論文では、アドホックネットワークにおけるチェックポイントプロトコルを示した。紛失メッセージとなる可能性のあるメッセージを、エンド-エンドではなく、ホップバイホップで検証し、メッセージログに保存する機構を導入することにより、各移動コンピュータが隣接移動コンピュータに対して状態情報とメッセージログを含むチェックポイント要求メッセージを一度だけ送信することにより、同期オーバーヘッドを削減することができる。

今後は、ネットワークトポロジの変化が大きくなり、ネットワークの分割が発生し各ネットワークでチェックポイントプロトコルが実行された後合流する場合になどおけるチェックポイントの設定手法について検討していく。

なお、本研究は東京電機大学総合研究所 研究課題 Q05J-05 として行なわれた。

#### 参考文献

- [1] "Radio Equipment and Systems (RES); HIPERLAN," ETSI, Functional Specifications (1995).
- [2] "Wireless (MAC) and (PHY) Specifications for Wireless Personal Area Networks," Standard IEEE 802.15.1 (2002).
- [3] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," Standard IEEE 802.11 (1999).
- [4] Acharya, A. and Badrinath, B.R., "Checkpointing Distributed Applications on Mobile Computers," Proc. of 3rd International Conference on Parallel and Distributed Information Systems, pp. 73-80 (1994).
- [5] Callaway, E.M., "Wireless Sensor Networks," Auerbach Publications (2003).
- [6] 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).
- [7] Corson, M.S. and Ephremides, A., "A Distributed Routing Algorithm for Mobile Wireless Networks," ACM Journal of Wireless Networks, vol. 1, No. 1, pp. 61-81 (1995).
- [8] Elnozahy, E.N. and Zwaenepoel, W., "On the use and Implementation of Message Logging," Proc. of the Fault-Tolerant Computing Symposium, pp. 298-307 (1994).
- [9] Gendelman, E., Bic, L.F. and Dillencourt, M.B., "An efficient checkpointing algorithm for distributed systems implementing reliable communication channels," Proc. of 18th International Symposium on Reliable Distributed Systems, pp. 290-291 (1999).
- [10] Higaki, H. and Takizawa, M., "Checkpoint-Recovery Protocol for Reliable Mobile Systems," Proc. of the 17th International Symposium on Reliable Distributed Systems, pp. 93-99 (1998).
- [11] Johnson, D.B., Maltz, D.A., Hu, Y.C., and Jetcheva, J.G., "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks," Internet Draft, draft-ietf-manet-dsr-04.txt (2000).
- [12] Koo, R. and Toueg, S., "Checkpointing and Rollback-Recovery for Distributed Systems," IEEE Trans. on Software Engineering, Vol. SE-13, No. 1, pp. 23-31 (1987).
- [13] Lamson, B.W., Paul, M. and Siegert, H.J., "Distributed Systems - Architecture and Implementation," Springer-Verlag, pp. 246-265 (1981).
- [14] Lesser, V., Ortiz, C.L. and Tambe, M., "Distributed Sensor Networks," Kluwer Academic Publications (2003).
- [15] Miyazaki, M., Morita, Y. and Higaki, H., "Hybrid Checkpoint Protocol for Mobile Networks with Unreliable Wireless Communication Channels," Proc. of the 2nd Asian International Mobile Computing Conference, pp. 164-171 (2002).
- [16] Morita, Y. and Higaki, H., "Checkpoint-Recovery for Mobile Computing Systems," Proc. of the 21st International Conference Distributed Computing Systems Workshops, pp. 479-484 (2001).
- [17] Neves, N. and Fuchs, W.K., "Adaptive Recovery for Mobile Environments," Communications of the ACM, Vol. 40, No. 1, pp. 69-74 (1997).
- [18] Perkins, C.E. and Royer, E.M., "Ad-hoc On-Demand Distance Vector Routing," Proceedings of IEEE Workshop on Mobile Computing Systems and Applications, pp. 99-100 (1999).
- [19] Strom, R. and Yemini, S., "Optimistic Recovery in distributed systems," ACM Trans. on Computer Systems, Vol. 3, No. 3, pp. 204-226 (1985).
- [20] Yao, B. and Fuchs, W.K., "Message logging optimization for wireless networks," Proc. of the 20th International Symposium on Reliable Distributed Systems, pp. 182-185 (2001).