

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7031424号

(P7031424)

(45)発行日 令和4年3月8日(2022.3.8)

(24)登録日 令和4年2月28日(2022.2.28)

(51)国際特許分類

F I

G 0 6 F 9/52 (2006.01)

G 0 6 F

9/52

1 5 0 Z

G 0 6 F 9/54 (2006.01)

G 0 6 F

9/54

B

請求項の数 10 (全27頁)

(21)出願番号 特願2018-58720(P2018-58720)
(22)出願日 平成30年3月26日(2018.3.26)
(65)公開番号 特開2019-169109(P2019-169109
A)
(43)公開日 令和1年10月3日(2019.10.3)
審査請求日 令和2年12月10日(2020.12.10)

(73)特許権者 000005223
富士通株式会社
神奈川県川崎市中原区上小田中4丁目1
番1号
(74)代理人 100104190
弁理士 酒井 昭徳
(72)発明者 高橋 英一
神奈川県川崎市中原区上小田中4丁目1
番1号 富士通株式会社内
(72)発明者 司波 章
神奈川県川崎市中原区上小田中4丁目1
番1号 富士通株式会社内
(72)発明者 松井 一樹
神奈川県川崎市中原区上小田中4丁目1
番1号 富士通株式会社内

最終頁に続く

(54)【発明の名称】 情報処理装置、情報処理システムおよび情報処理プログラム

(57)【特許請求の範囲】

【請求項1】

複数のマイクロサービス間でメッセージ通信し、それぞれの前記マイクロサービスが到着順序でメッセージ処理する情報処理装置において、
受信したメッセージを到着順序で格納するキューと、
前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、
受信した複数の前記メッセージの到着順序と複数の前記メッセージのタイムスタンプとに基づき、複数の前記メッセージのうち遅延して到着した前記メッセージがあれば正しい到着順序にして複数の前記メッセージを処理し、前記メッセージを受信する下流の情報処理装置に対し、正しい到着順序で前記メッセージを処理させる修正通知を送信するメッセージ制御部と、を有し、
前記メッセージ制御部は、
前記サービス処理部のメッセージ処理毎の処理状態をスナップショットとして処理順に保持する履歴記録部と、
前記キューに格納された前記メッセージと、前記履歴記録部に保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信する検査部と、
前記修正通知の受信時に、前記キューおよび前記サービス処理部に対し、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこなうロールバック制御部と、

を有し、
前記検査部は、
下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信することを特徴とする情報処理装置。

【請求項 2】

前記ロールバック制御部は、
前記修正通知の受信時に、当該修正通知に基づき、前記キューに対し、正しく修正した後のメッセージ列を格納し、前記サービス処理部に対し、遅延して到着した前記メッセージ以前のメッセージ処理状態にロールバックさせることを特徴とする請求項 1 に記載の情報処理装置。

10

【請求項 3】

前記ロールバック制御部は、
前記修正通知の受信時に、当該修正通知に基づき、前記サービス処理部に対し、到着順序が正しくない前記メッセージのメッセージ処理をキャンセルさせることを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

前記ロールバック制御部は、
前記修正通知の受信時に、前記キューに格納された修正対象の前記メッセージを削除することを特徴とする請求項 2 または 3 に記載の情報処理装置。

20

【請求項 5】

上流の複数の情報処理装置から合流して前記メッセージを受信するシステム構成の場合、前記ロールバック制御部は、
前記修正通知の受信時に、前記サービス処理部によりメッセージ処理を停止し、前記キューの先頭のメッセージが正しく修正されたメッセージであれば前記停止を解除することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 6】

前記履歴記録部は、
下流の情報処理装置の前記サービス処理部の数と、キューサイズに基づく記録容量を有することを特徴とする請求項 1 ～ 5 のいずれか一つに記載の情報処理装置。

30

【請求項 7】

複数のマイクロサービス間でメッセージ通信し、それぞれの前記マイクロサービスが到着順序でメッセージ処理する複数の情報処理装置を含む情報処理システムにおいて、前記情報処理装置は、それぞれ、
受信したメッセージを到着順序で格納するキューと、
前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、
受信した複数の前記メッセージの到着順序と複数の前記メッセージのタイムスタンプとに基づき、複数の前記メッセージのうち遅延して到着した前記メッセージがあれば正しい到着順序にして複数の前記メッセージを処理し、前記メッセージを受信する下流の情報処理装置に対し、正しい到着順序で前記メッセージを処理させる修正通知を送信するメッセージ制御部と、を有し、
前記メッセージ制御部は、
前記サービス処理部のメッセージ処理毎の処理状態をスナップショットとして処理順に保持する履歴記録部と、
前記キューに格納された前記メッセージと、前記履歴記録部に保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信する検査部と、
前記修正通知の受信時に、前記キューおよび前記サービス処理部に対し、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこなうロールバック制御部と、

40

50

を有し、
前記検査部は、
下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信することを特徴とする情報処理システム。

【請求項 8】

複数のマイクロサービス間でメッセージ通信し、それぞれの前記マイクロサービスが到着順序でメッセージ処理する複数の情報処置装置と、メッセージ制御装置とを含む情報処理システムにおいて、

前記情報処理装置は、それぞれ、

受信したメッセージを到着順序で格納するキューと、

前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、を有し、

前記メッセージ制御装置は、

前記情報処理装置が受信した複数の前記メッセージの到着順序と複数の前記メッセージのタイムスタンプとに基づき、複数の前記メッセージのうち遅延して到着した前記メッセージがあれば正しい到着順序にして複数の前記メッセージを処理し、前記メッセージを受信する下流の情報処理装置に対し、正しい到着順序で前記メッセージを処理させる修正通知を送信するメッセージ制御部を有し、

前記メッセージ制御部は、

前記サービス処理部のメッセージ処理毎の処理状態をスナップショットとして処理順に保持する履歴記録部と、

前記キューに格納された前記メッセージと、前記履歴記録部に保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信する検査部と、

前記修正通知の受信時に、前記キューおよび前記サービス処理部に対し、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこなうロールバック制御部と、を有し、

前記検査部は、

下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信することを特徴とする情報処理システム。

【請求項 9】

複数のマイクロサービス間でメッセージ通信し、それぞれの前記マイクロサービスが到着順序でメッセージ処理する情報処置装置のコンピュータに、

受信したメッセージを到着順序でキューに格納し、

前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこない、受信した複数の前記メッセージの到着順序と複数の前記メッセージのタイムスタンプとに基づき、複数の前記メッセージのうち遅延して到着した前記メッセージがあれば正しい到着順序にして複数の前記メッセージを処理し、

前記メッセージを受信する下流の情報処理装置に対し、正しい到着順序で前記メッセージを処理させる修正通知を送信し、

さらに、前記メッセージ処理毎の処理状態をスナップショットとして処理順に保持し、

前記キューに格納された前記メッセージと、保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信し、

前記修正通知の受信時に、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこない、

下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信

10

20

30

40

50

する、

処理を実行させることを特徴とする情報処理プログラム。

【請求項 10】

前記修正通知の受信時に、当該修正通知に基づき、正しく修正した後のメッセージ列をキューに格納し、

前記マイクロサービスの処理をおこなうサービス処理部に対し、遅延して到着した前記メッセージ以前のメッセージ処理状態にロールバックさせる、

処理を実行させることを特徴とする請求項 9 に記載の情報処理プログラム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、情報処理装置、情報処理システムおよび情報処理プログラムに関する。

【背景技術】

【0002】

主にクラウド上の分散処理システムでは、マイクロサービスアーキテクチャと呼ぶ分散型システムを構築し、マイクロサービスを分散処理する。マイクロサービスの多くは検索や認証など単機能の小さなマイクロサービスであり、複数のマイクロサービスを組合せてマイクロサービスを構築する。たとえば、Web インタフェースマイクロサービス、認証マイクロサービス、検索マイクロサービス、ショッピングカートマイクロサービス、データベースマイクロサービスの 5 つのマイクロサービスを組合せてネット販売マイクロサービスを構築する例がある。

20

【0003】

マイクロサービスを組合せる方法としてメッセージングが広く用いられ、メッセージングによりマイクロサービス同士をメッセージのやり取りで連携させることができる。メッセージングでは、メッセージ送信側（上流）のマイクロサービスは、受信側（下流）のマイクロサービスでのメッセージ受信を待たずに、つぎのメッセージ処理をおこなう非同期実行に適している。

【0004】

上流のマイクロサービスは、下流のマイクロサービスがメッセージ処理するか否かに関わらず、つぎのメッセージを順次送信し、下流のマイクロサービスは、メッセージをキューと呼ぶバッファに格納する。下流のマイクロサービスは、キューからメッセージを取り出し内容に応じたメッセージ処理をおこなうことを繰り返す。

30

【0005】

従来技術として、第 1 の処理から第 2 の処理へ送信するメッセージの内容が一致するように、タイムスタンプを用いロールバックさせることで、並行処理を同期化させる技術が開示されている（たとえば、下記特許文献 1 参照。）。また、各応用プログラムが送受信ログを収集して半順序関係を示すタイムスタンプを付与し、応用プログラム間でロールバック可能な最新のチェックポイントを探索して、応用プログラムの回復をおこなう技術が開示されている（たとえば、下記特許文献 2 参照。）。

【先行技術文献】

40

【特許文献】

【0006】

【文献】特開平 6 - 028199 号公報

特開平 7 - 271620 号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、従来技術では、マイクロサービス間でおこなうメッセージ通信のメッセージ順序を補償できなかった。

【0008】

50

メッセージングでは、様々な原因でメッセージが送信時と異なる到着順序で下流に到着する場合がある。たとえば、メッセージが後発メッセージと異なる遅いネットワーク経路で伝送され、ネットワーク上でこのメッセージがロストすると、後発メッセージ送信後にメッセージが再送される。また、メッセージサイズが大きいため分割送信されると、分割した全てが下流に到着する前に後発のより小さいメッセージが先に下流に到着してしまう。

【0009】

メッセージの到着順序が送信順序と異なると正しくマイクロサービス処理できない場合がある。たとえば、センサや端末からのデータ（メッセージ）を扱うマイクロサービスは、メッセージの到着順序が変わると間違った処理をおこなう問題が生じる。たとえば、実世界で起きる事象をメッセージとして取り扱うマイクロサービスにおいては、事象に対応したメッセージを事象が起きた際の送信順序で処理する必要がある。しかし、メッセージの到着が遅れる原因は多様であり、伝送路で保証しようとすると、専用のネットワークを持つ高価なシステムが必要になってしまう。

【0010】

また、従来のロールバック技術では、マイクロサービス状態の保存（スナップショット）をどこまで（何世代）保存すればよいかを定めることができず、運用者が経験上十分と考える任意の世代数に定めている。この場合、下流では余裕を持った世代数分のスナップショットを保存するため必要以上のスナップショットを保存することになり、主に記憶資源（ハードウェア）が高コスト化する問題が生じる。

【0011】

ここで、送信側（上流）と受信側（下流）が1：1接続のマイクロサービス間ではMQ（メッセージキューイング）などの順序性保証機構によりメッセージの到着順序を補償する方法がある。しかし、マイクロサービス間の一部接続がn：1のシステムでは下流でのメッセージの到着順序を補償できない。

【0012】

一つの側面では、本発明は、メッセージ到着遅れが生じても複数のマイクロサービスがそれぞれ正しいメッセージ到着順序でメッセージ処理をおこなうことを目的とする。

【課題を解決するための手段】

【0013】

一つの実施の形態では、複数のマイクロサービス間でメッセージ通信し、それぞれの前記マイクロサービスが到着順序でメッセージ処理する情報処理装置において、受信したメッセージを到着順序で格納するキューと、前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、受信した複数の前記メッセージの到着順序と複数の前記メッセージのタイムスタンプとに基づき、複数の前記メッセージのうち遅延して到着した前記メッセージがあれば正しい到着順序にして複数の前記メッセージを処理し、前記メッセージを受信する下流の情報処理装置に対し、正しい到着順序で前記メッセージを処理させる修正通知を送信するメッセージ制御部と、を有し、前記メッセージ制御部は、前記サービス処理部のメッセージ処理毎の処理状態をスナップショットとして処理順に保持する履歴記録部と、前記キューに格納された前記メッセージと、前記履歴記録部に保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信する検査部と、前記修正通知の受信時に、前記キューおよび前記サービス処理部に対し、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこなうロールバック制御部と、を有し、前記検査部は、下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信する情報処理装置、情報処理システムおよび情報処理プログラムが提供される。

【発明の効果】

【0014】

本発明の一側面によれば、メッセージ到着遅れが生じても複数のマイクロサービスがそれ

10

20

30

40

50

ぞれ正しいメッセージ到着順序でメッセージ処理をおこなうことができる。

【図面の簡単な説明】

【0015】

【図1】図1は、実施の形態にかかるシステムの全体構成の一例を示す説明図である。

【図2A】図2Aは、実施の形態にかかる情報処理装置の機能的構成の一例を示す説明図である。

【図2B】図2Bは、実施の形態にかかる情報処理装置間の接続構成の一例を示す説明図である。

【図3】図3は、実施の形態にかかる情報処理装置のハードウェア構成の一例を示すブロック図である。

【図4】図4は、実施の形態にかかる情報処理装置がおこなう全体の処理の手順の一例を示すフローチャートである。

【図5】図5は、実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャート（その1）である。

【図6】図6は、実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャート（その2）である。

【図7】図7は、実施の形態にかかる情報処理装置の停止中処理の手順の一例を示すフローチャートである。

【図8】図8は、実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その1）である。

【図9】図9は、実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その2）である。

【図10】図10は、実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その3）である。

【図11】図11は、実施の形態にかかる情報処理装置に必要なスナップショットのサイズを示す説明図である。

【図12】図12は、従来技術によるメッセージ通信の例を示す説明図である。

【図13】図13は、 $n:1$ のメッセージ通信で順序保証が必要な事象を示す説明図である。

【図14A】図14Aは、従来技術による複数のマイクロサービスを連携させたときの問題を示す説明図（その1）である。

【図14B】図14Bは、従来技術による複数のマイクロサービスを連携させたときの問題を示す説明図（その2）である。

【発明を実施するための形態】

【0016】

以下に図面を参照して、開示の情報処理装置、情報処理システムおよび情報処理プログラムの実施の形態を詳細に説明する。

【0017】

図1は、実施の形態にかかるシステムの全体構成の一例を示す説明図である。図1を用いて実施の形態の情報処理システムがおこなうメッセージの到着順序性の保証の概要を説明する。

【0018】

実施の形態の情報処理システム100は、受信した要求マイクロサービスに対し、マイクロサービスA～E毎のマイクロサービス処理部101（101a～101e）がメッセージ（データ）に対し、それぞれ異なるメッセージ処理（データ処理）を分散しておこなう。

【0019】

要求マイクロサービスとは、情報処理システム100外部のクライアント、たとえば、ユーザの端末操作による所定のマイクロサービスの要求、あるいはセンサ装置等の検出データに基づくデータ処理の要求等である。

【0020】

10

20

30

40

50

この際、図 1 の矢印 (→) に示すように、複数のマイクロサービス処理部 101 は、それぞれ上流から入力されたメッセージをメッセージ処理した後、下流のマイクロサービス処理部 101 に送信する。

【0021】

各マイクロサービス処理部 101 は、たとえば、クラウド上に配置された複数のサーバがネットワーク接続により相互にメッセージを送受信することができる。以下の説明では、上流、下流は、メッセージが流れる方向を示し、送信側のマイクロサービス処理部 101 が上流に位置してサービス処理の結果をメッセージとして送信し、受信側のマイクロサービス処理部 101 が下流に位置してメッセージを受信する。

【0022】

図 1 に示す例では、マイクロサービス A1、A2 をそれぞれ並列しておこなうマイクロサービス処理部 101a を有する。このような構成においては、マイクロサービス A1、A2 のマイクロサービス処理部 101a は、順序保証をおこなわずにメッセージを送信する。

【0023】

実施の形態では、メッセージ制御部 102 を設け、複数のマイクロサービス処理部 101 (101a~101e) でのメッセージの到着順序が正しくなるよう、到着順序性の制御 (ロールバック制御) をおこなう。メッセージ制御部 102 は、図 1 に示すように、マイクロサービス処理部 101 と独立して配置してもよい。この場合、メッセージ制御部 102 は、サービス処理部 101 と異なる他の機器、たとえば、メッセージ・ブローカー (サーバ等を用いたメッセージ制御装置) として独立して配置することができる。また、後述するが、メッセージ制御部 102 は、各マイクロサービス処理部 101 と同一の機器内 (たとえばサーバ) に配置してもよい。

【0024】

上記情報処理システム 100 には、最初の発生時を示すタイムスタンプを有するメッセージが入力される。そして、上流からメッセージを受信したマイクロサービス処理部 101 は、メッセージ処理毎のマイクロサービスのサービス状態のスナップショットを履歴として履歴記録部 103 に記録保持する。スナップショットはマイクロサービス処理終了時のタイムスタンプを含む。

【0025】

メッセージ制御部 102 は、キュー 201 (図 2 参照) に入ったメッセージのタイムスタンプを参照し、キュー 201 内のメッセージのタイムスタンプ (時刻) を比較することで、メッセージの到着遅れを判断する。

【0026】

また、メッセージ制御部 102 は、各マイクロサービス処理部 101 が受信したメッセージのタイムスタンプと、最新状態のスナップショットのタイムスタンプとを比較して、各マイクロサービス処理部 101 におけるメッセージの到着遅れの有無を判断する。

【0027】

そして、メッセージ制御部 102 は、メッセージの到着遅れが生じたマイクロサービス処理部 101 および下流の全マイクロサービス処理部 101 に対し、遅れたメッセージ以前のメッセージ処理状態へロールバックする制御をおこなう。これにより、メッセージの到着遅れが生じたマイクロサービス処理部 101 および下流の全マイクロサービス処理部 101 は、正しいメッセージの到着順序でマイクロサービスを再実行する。これにより、要求メッセージに対して、情報処理システム 100 内でメッセージの到着遅れが生じて、全てのメッセージ制御部 102 が順序保証してメッセージ処理できるようになる。

【0028】

図 2A は、実施の形態にかかる情報処理装置の機能的構成の一例を示す説明図である。図 1 に示した一つのマイクロサービスのマイクロサービス処理部 101 は、要求マイクロサービスのうちの一つのマイクロサービス (データ処理) をおこなう一つの情報処理装置 200 に配置される。

【0029】

10

20

30

40

50

図2Aの(a)は、一つの情報処理装置200(マイクロサービス処理部101)毎にメッセージ制御部102を配置した構成例の機能ブロック図である。また、図2Aの(b)には、情報処理装置200内で保持する情報例(メッセージ、データ等)の図表を示す。

【0030】

図2Aの(a)に示すように、情報処理装置200は、上述したマイクロサービス処理部101と、メッセージ制御部102と、キュー201とを含む。また、メッセージ制御部102は、履歴記録部103と、検査部212と、ロールバック制御部213とを含む。

【0031】

キュー201は、上流から受信したメッセージ(データ)を到着順に保持し、マイクロサービス処理部101に出力する。メッセージmiは、内容(データ)と、メッセージの識別子(ID)と、生成時刻(タイムスタンプ:ST)を含む(図2Aの(b)参照)。

【0032】

マイクロサービス処理部101は、キュー201から出力されたメッセージ(データ)に対し、所定のマイクロサービス機能に基づくメッセージ処理(データ処理)をおこない、処理後のメッセージを下流の情報処理装置200に送信する。

【0033】

履歴記録部103は、マイクロサービス処理部101によるメッセージ処理毎の処理メッセージと、スナップショットを順次記録保持する。たとえば、履歴記録部103は、メッセージ処理順にスナップショットを図2Aの(b)のように下から上に順次蓄積記録していく。スナップショットはマイクロサービス処理部101によるマイクロサービス処理終了時のタイムスタンプを含む。

【0034】

検査部212は、キュー201が受信したメッセージの到着遅れの有無を検出する。検査部212は、履歴記録部103で処理したメッセージ処理時のタイムスタンプを参照し、キュー201の該当するメッセージのタイムスタンプ(生成時刻ST)が新しいか、あるいは古いかに基づき、メッセージの到着遅れの有無を検出する。

【0035】

そして、検査部212は、受信したメッセージのうち、どのメッセージの到着順序が正しくないか、たとえば到着遅れあり、を示す情報を修正通知として下流の情報処理装置200に送信する。

【0036】

修正通知は、修正対象メッセージIDリスト(CML)と、修正後先頭メッセージID(CTM)を含む(図2Aの(b)参照)。修正対象メッセージIDリスト(CML)は、順序変更の影響を受けるメッセージのIDをリスト化した集合である。修正後先頭メッセージID(CTM)は、到着順序を修正したメッセージ列の最初のメッセージのIDである。この修正通知は、メッセージの到着遅れを検出したメッセージ制御部102(検査部212)が生成および送信するものであり、上述したメッセージと同様に、上流の情報処理装置200から下流の情報処理装置200に向けて送信される。

【0037】

また、図1のように、マイクロサービス処理部101と独立してメッセージ制御部102を集約する形で配置した場合、メッセージ制御部102は、該当する下流の複数のマイクロサービス処理部101(各サーバ等)に一斉に修正通知をブロードキャストする。

【0038】

また、検査部212は、上流の情報処理装置200から修正通知を受信した場合、キュー201に対し、まだ処理中のメッセージが残っている場合、キュー201内に残っているメッセージを削除(データ消去等)する制御をおこなう。また、検査部212は、メッセージ削除に対応してキュー201内のメッセージのソートをおこなう。

【0039】

ロールバック制御部213は、上流の情報処理装置200から修正通知を受信した場合、マイクロサービス処理部101に対するメッセージ処理をキャンセルし、修正対象のメッ

10

20

30

40

50

セージ処理前まで状態を戻すロールバック制御をおこなう。このロールバック制御では、修正通知の受信により、履歴記録部 103 が記録保持するスナップショットのうち、メッセージ到着遅れを起こしたメッセージをキュー 201 に戻す処理をおこなう。

【0040】

図 2 B は、実施の形態にかかる情報処理装置間の接続構成の一例を示す説明図である。図 2 B には、複数のマイクロサービス間で接続した複数の情報処理装置 200 A、200 B を示す。各情報処理装置 200 A、200 B の内部構成は、図 2 A (a) と同様である。図 2 B に示すように、メッセージおよび修正通知は、上流の情報処理装置 200 A から下流の情報処理装置 200 B に伝播される。

【0041】

図 3 は、実施の形態にかかる情報処理装置のハードウェア構成の一例を示すブロック図である。図 2 A に示した情報処理装置 200 は、図 3 に示すハードウェアを用いることができる。

【0042】

たとえば、情報処理装置 200 は、CPU (Central Processing Unit) 301、メモリ 302、ネットワークインタフェース (I/F) 303、記録媒体 I/F 304、記録媒体 305、を含む。300 は各部を接続するバスである。

【0043】

CPU 301 は、情報処理装置 200 の全体の制御を司る制御部として機能する演算処理装置である。メモリ 302 は、不揮発性メモリおよび揮発性メモリを含む。不揮発性メモリは、たとえば、CPU 301 のプログラムを格納する ROM (Read Only Memory) である。揮発性メモリは、たとえば、CPU 301 のワークエリアとして使用される DRAM (Dynamic Random Access Memory)、SRAM (Static Random Access Memory) 等である。

【0044】

ネットワーク I/F 303 は、ネットワーク 330 を通じて他の情報処理装置 200 との間で処理に必要な情報 (上記メッセージや修正通知等) を送受信する。ネットワーク I/F 303 は、LAN (Local Area Network)、WAN (Wide Area Network)、インターネットなどのネットワーク 330 を介して他の情報処理装置 200 に通信接続される。

【0045】

記録媒体 I/F 304 は、CPU 301 が処理した情報を記録媒体 305 との間で読み書きするためのインタフェースである。記録媒体 305 は、メモリ 302 を補助する記録装置である。たとえば、HDD (Hard Disk Drive)、SSD (Solid State Drive)、USB (Universal Serial Bus) フラッシュドライブ等を用いることができる。

【0046】

メモリ 302 または記録媒体 305 に記録されたプログラムを CPU 301 が実行することにより、図 2 A に示した情報処理装置 200 のマイクロサービス処理部 101 およびメッセージ制御部 102 (検査部 212、ロールバック制御部 213) の機能を実現する。また、メモリ 302 や記録媒体 305 を用いて、図 2 A に示したキュー 201、履歴記録部 103 の情報を保持することができる。

【0047】

また、図 3 に示したハードウェア構成は、図 1 のシステム構成例の場合、各マイクロサービスを実行処理するマイクロサービス処理部 101 の機能を有する所定のサーバ、メッセージ制御部 102 の機能を有するサーバ、にそれぞれ同様に適用することができる。これらの各サーバにおいても、CPU 301 を用いてプログラム実行する場合、CPU 301 がメモリ 302 や記録媒体 305 から読み出したプログラムを実行することにより、それぞれのサーバの機能を実現することができる。

【0048】

10

20

30

40

50

図 4 は、実施の形態にかかる情報処理装置がおこなう全体の処理の手順の一例を示すフローチャートである。図 4 の処理は、一つの情報処理装置 200 に設けられるメッセージ制御部 102 が主におこなう。

【0049】

情報処理装置 200 において、キュー 201 に受信したメッセージ m が時間順を有して入る（ステップ S401）。キュー 201 は、メッセージ m をタイムスタンプの順番に保持する。

【0050】

そして、メッセージ制御部 102 は、マイクロサービス処理部 101 のマイクロサービスが実行状態であるか、あるいは停止状態であるかを判断する（ステップ S402）。マイクロサービスが実行状態であれば（ステップ S402：実行状態）、マイクロサービス実行中に対応した実行中処理を実施する（ステップ S403）。一方、マイクロサービスが停止状態であれば（ステップ S402：停止状態）、マイクロサービス停止中に対応した停止中処理を実施する（ステップ S404）。ステップ S403 あるいはステップ S404 の処理実行により、マイクロサービス処理部 101 は以上の処理を終了する。

【0051】

これらマイクロサービスの実行状態、および停止状態については後述する。通常は、マイクロサービス処理部 101 によるマイクロサービス（メッセージ処理）が継続し、実行状態となっている（後述する問題ケース 1～3 の処理時に相当）。一方、上流からの修正通知に基づきマイクロサービス処理部 101 によるマイクロサービスを停止（stop）させる場合、停止状態となる（後述する問題ケース 4 の処理時に相当）。

【0052】

図 5 は、実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャートである。図 5 の処理は、図 4 のステップ S403 の処理の詳細であり、上流に配置されてメッセージ到着遅れを検出した情報処理装置 200 に設けられるメッセージ制御部 102 がおこなう。

【0053】

はじめに、メッセージ制御部 102 の検査部 212 は、履歴記録部 103 に記録保持されているスナップショット（履歴記録） $SS[i]$ （ $i=1, \dots, n$ ）を調べる。そして、検査部 212 は、マイクロサービス処理部 101 がメッセージ処理したメッセージの中でキュー 201 の先頭メッセージ m よりもタイムスタンプが後のメッセージの集合 MM を求める（ステップ S501）。この MM は、遅延等が生じた修正対象メッセージ集合である。

【0054】

そして、検査部 212 は、 MM が空であるか否かを判断する（ステップ S502）。 MM が空であれば（ステップ S502：Yes）、ステップ S507 の処理に移行する。

【0055】

一方、 MM が空でなければ（ステップ S502：No）、検査部 212 は、メッセージ m が到着遅れと判定する。そして、検査部 212 は、下流に位置する全てのマイクロサービスに対応した各情報処理装置 200 のメッセージ制御部 102（ロールバック制御部 213）へ修正通知（CML および CTM）を送信する（ステップ S503）。

【0056】

修正対象メッセージ ID 集合 CML は、上記 MM の要素のメッセージ ID 集合であり、修正後先頭メッセージ ID（CTM）は、メッセージ m の ID である。

【0057】

つぎに、メッセージ制御部 102 のロールバック制御部 213 は、履歴記録部 103 に記録保持されているスナップショット $SS[i]$ （ $i=1, \dots, n$ ）から、ロールバック状態 $SS[r]$ を求める（ステップ S504）。ここで、 m_j はメッセージ m のタイムスタンプよりも後のメッセージで、かつ、スナップショット SS の中で最もタイムスタンプが早いメッセージである。そして、この m_j をマイクロサービス処理部 101 がメッセージ処理したときのスナップショット $SS[r+1]$ を求め、 $SS[r+1]$ の直前のスナッ

10

20

30

40

50

プショット $SS[r]$ のロールバック状態を求める。

【0058】

つぎに、ロールバック制御部 213 は、マイクロサービス処理部 101 のマイクロサービス状態を $SS[r]$ に戻し、履歴記録部 103 が記録保持するスナップショットのうち、スナップショット $SS[r+1] \sim SS[n]$ を破棄する。また、変数 $n=r$ とする（ステップ S505）。

【0059】

つぎに、ロールバック制御部 213 は、キュー 201 にメッセージ m と MM の各メッセージを時間順（タイムスタンプ順）でセットする（ステップ S506）。そして、マイクロサービス処理部 101 は、キュー 201 からメッセージ m を取り出しマイクロサービスを実行する（ステップ S507）。この後、履歴記録部 103 は、その時点のマイクロサービスの状態とメッセージ m をスナップショット $SS[n+1]$ として追加記録し（ステップ S508）、以上の処理を終了する。なお、履歴記録部 103 のスナップショット記録にかかる記録容量（最大サイズ）については後述する。

【0060】

図 6 は、実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャートである。図 6 の処理は、図 4 のステップ S403 の処理の詳細であり、下流に配置され修正通知を受信した情報処理装置 200 に設けられるメッセージ制御部 102（ロールバック制御部 213）がおこなう。

【0061】

ロールバック制御部 213 は、上流の情報処理装置 200 のメッセージ制御部 102 から修正通知を受信する（ステップ S601）。この修正通知の受信時、ロールバック制御部 213 は、履歴記録部 103 に記録保持されているスナップショット $SS[i]$ （ $i=1, \dots, n$ ）から、ロールバック状態 $SS[r]$ を求める（ステップ S602）。 $SS[r]$ はマイクロサービス処理部 101 がメッセージ処理したメッセージ ID が修正通知の CML にあるスナップショットの中で、処理したメッセージのタイムスタンプが最も古いスナップショット $SS[r+1]$ の直前のスナップショット $SS[r]$ である。

【0062】

つぎに、ロールバック制御部 213 は、履歴記録部 103 内の $SS[r]$ の有無を判断する（ステップ S603）。履歴記録部 103 に $SS[r]$ がない場合（ステップ S603：No）、以降の処理をおこなわずに処理終了する。これは、メッセージ到着遅れを起こしたメッセージよりも古いメッセージ処理しか実行していない状態であり、ロールバックは不要である。

【0063】

一方、ロールバック制御部 213 は、履歴記録部 103 に $SS[r]$ がある場合（ステップ S603：Yes）、メッセージ到着遅れを起こしたメッセージの後続メッセージをマイクロサービス処理部 101 で先にメッセージ処理してしまった状態を意味する。この場合、ロールバック制御部 213 は、履歴記録部 103 のスナップショット $SS[j]$ （ $j=1, \dots, n$ ）の処理メッセージの中で、修正通知の CML に ID が含まれないメッセージを時間順（タイムスタンプ順）にキュー 201 へ戻す（ステップ S604）。

【0064】

そして、ロールバック制御部 213 は、マイクロサービス状態を $SS[r]$ に戻す。また、履歴記録部 103 に記録保持されている $SS[r]$ 以後のスナップショット $SS[r+1] \sim SS[n]$ を破棄する。また、 $n=r$ とする（ステップ S605）。

【0065】

この後、ロールバック制御部 213 は、マイクロサービスを停止状態にし（ステップ S606）、以上の処理を終了する。この後、マイクロサービスが停止状態となった情報処理装置 200 のメッセージ制御部 102 は、図 4 に示した停止中処理（ステップ S404）を実行する。

【0066】

上述したように、マイクロサービス処理部 101 が実行するマイクロサービスには、実行状態と停止状態の 2 つの状態がある。実行状態では上述した図 5 のステップ S 507 の処理でメッセージをキュー 201 から取り出して実行する。一方、ロールバック制御部 213 が修正通知を受信すると図 6 のステップ S 606 でマイクロサービスが停止状態になる。この停止状態の間中は、マイクロサービス処理部 101 のマイクロサービス（メッセージ処理）の動作は停止（stop）し、キュー 201 からのメッセージ取得をおこなわない。

【0067】

図 7 は、実施の形態にかかる情報処理装置の停止中処理の手順の一例を示すフローチャートである。図 7 の処理は、図 4 のステップ S 404 の処理の詳細であり、下流に配置され修正通知を受信した情報処理装置 200 に設けられるメッセージ制御部 102 がおこなう。

10

【0068】

メッセージ制御部 102 は、マイクロサービス停止中にメッセージを受信した場合、図 4 に示したステップ S 404 の停止中処理を実行する。この停止中処理では、はじめにメッセージ制御部 102 の検査部 212 がキュー 201 の先頭メッセージの ID（IDt）と、修正通知の修正後先頭メッセージ ID（CTM）とを比較する（ステップ S 701）。

【0069】

この後、検査部 212 は、先頭メッセージの ID（IDt）と、修正通知の修正後先頭メッセージ ID（CTM）とが同じであるか否かを判断する（ステップ S 702）。判断の結果、IDt が CTM と異なっていれば（ステップ S 702：No）、再実行するメッセージが到着していないと判断し、以後の処理をおこなわずに処理終了する。

20

【0070】

一方、判断の結果、IDt が CTM と同一であれば（ステップ S 702：Yes）、再実行するメッセージである。このため、メッセージ制御部 102（検査部 212）は、マイクロサービスを実行状態へ変え、マイクロサービス処理部 101 に対し、キュー 201 からメッセージを取り出し、マイクロサービスを実行させる（ステップ S 703）。

【0071】

図 8～図 10 は、実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図である。これらの図を用いて情報処理装置 200 におけるメッセージの到着順序が異なる各問題ケースに対応した修正処理の概要を説明する。これら図 8～図 10 には、所定のマイクロサービスを実行する一つの情報処理装置 200 あたり、一つのキュー 201 とマイクロサービス処理部 101 の組として記載してある。

30

【0072】

図 8 には、（a）メッセージの到着順序が正常な場合の処理例、（b）問題ケース 1 に対する修正処理例 1、（c）問題ケース 2 に対する修正処理例 2 を示す。

【0073】

図 8（a）メッセージの到着順序が正常な場合、（1）メッセージ m1、m2 の順序で受信すると、まずメッセージ m1 がキュー 201 に入る。この後、（2）マイクロサービス処理部 101 がメッセージ m1 をメッセージ処理し、この際、メッセージ m2 がキュー 201 に入る。この後、（3）マイクロサービス処理部 101 はメッセージ処理後のメッセージ m1 を送信し、メッセージ m2 をメッセージ処理する。この際、キュー 201 は空となる。

40

【0074】

図 8（b）問題ケース 1 は、一つの情報処理装置 200 内のみでメッセージの到着順序に遅れが生じた状態を示す。本来は図 8（a）に示すように、メッセージ m1、m2 の順序に受信するが、メッセージ m1 に遅延が生じたとする。そして（1）メッセージ m2、m1 の順序で受信すると、メッセージ m2 が先にキュー 201 に入る。この後、（2）マイクロサービス処理部 101 がメッセージ m2 をメッセージ処理し、メッセージ m1 がキュー 201 に入る。

【0075】

50

上述したように、メッセージm1、m2はタイムスタンプ（生成時刻）を持つ。そして、実施の形態では、情報処理装置200（メッセージ制御部102）がタイムスタンプによりメッセージm1の到着遅れ（順序の逆転）を判断する。そして、メッセージ制御部102のロールバックにより、メッセージを正常な到着順序m1、m2に修正する。

【0076】

これにより、（3）キュー201には、メッセージm1、m2の順序で入り、マイクロサービス処理部101はメッセージm2のメッセージ処理をキャンセルする。その後、（4）正常なメッセージの到着順序m1、m2でメッセージ処理することができる。この問題ケース1のように、一つの情報処理装置200内でのメッセージの到着遅れの影響は、タイムスタンプを用いた比較のみで修正処理できる。

【0077】

図8（c）問題ケース2では、上流の情報処理装置200Aと、下流の情報処理装置200Bのキュー201およびマイクロサービス処理部101を記載し、メッセージm1の遅延が下流の情報処理装置200Bにまで及ぶケースを示す。

【0078】

はじめに、（1）メッセージm2、m1の順序を有し、メッセージm2は上流の情報処理装置200Aではメッセージ処理済みであり、下流の情報処理装置200Bのキュー201に入り、メッセージm1が上流の情報処理装置200Aで未受信とする。

【0079】

その後、（2）上流の情報処理装置200Aのキュー201にメッセージm1が入り、下流の情報処理装置200Bはマイクロサービス処理部101がメッセージm2をメッセージ処理したとする。

【0080】

実施の形態では、上流の情報処理装置200Aのマイクロサービス処理部101は、メッセージ処理時のスナップショットm2を記録保持している。情報処理装置200Aのメッセージ制御部102は、キュー201のメッセージm1のタイムスタンプと、スナップショットm2の比較により、メッセージm1の到着遅れ（順序の逆転）を判断する。そして、情報処理装置200A、200Bでは、メッセージ制御部102がm2処理前までロールバックすることにより、メッセージを正常な到着順序m1、m2に修正する。

【0081】

この際、メッセージの到着順序が正常でないと判断した情報処理装置200Aは、下流の情報処理装置200Bに対して上述した修正通知を送信する。

【0082】

これにより、（3）上流の情報処理装置200Aのキュー201には、メッセージm1、m2の順序で入る。また、下流の情報処理装置200Bのマイクロサービス処理部101はメッセージm2のメッセージ処理をキャンセルする。その後、（4）上流の情報処理装置200Aでは、正常なメッセージの到着順序m1、m2でメッセージ処理することができる。なお、下流の情報処理装置200Bでも、上流の情報処理装置200Aからメッセージm1、m2の順序にメッセージを受信し、正常にメッセージ処理できる。

【0083】

図9には、問題ケース3に対する修正処理例3を示す。この問題ケース3では、正常時の到着順序がメッセージm1、m2、m3、m4、m5、m6であるが、メッセージm1が遅延したため、到着順序がm2、m3、m4、m5、m6、m1の順序に入れ替わった状態を示す。

【0084】

はじめに、（1）メッセージm2、m3、m4、m5、m6、m1のうち、メッセージm2は最下流の情報処理装置200Cのキュー201に入ったとする。メッセージm3は中間の情報処理装置200Bのマイクロサービス処理部101でメッセージ処理し、メッセージm4が情報処理装置200Bのキュー201に入ったとする。また、メッセージm5は最上流の情報処理装置200Aのマイクロサービス処理部101でメッセージ処理し、

10

20

30

40

50

メッセージm6が情報処理装置200Aのキュー201に入ったとする。また、メッセージm1は情報処理装置200Aで未受信とする。

【0085】

この場合、実施の形態では、最上流の情報処理装置200Aでは、メッセージm2、m3、m4、m5のスナップショットを記録保持している。中間の情報処理装置200Bでは、メッセージm2、m3のスナップショットを記録保持している。

【0086】

そして、(2) 情報処理装置200Aのメッセージ制御部102は、キュー201のメッセージm1のタイムスタンプと、スナップショットm2～m6とを比較する。この比較により、メッセージm1の到着遅れ(順序の逆転)を判断し、また、メッセージm2～m6が順序の修正対象であると判断する。そして、情報処理装置200Aのメッセージ制御部102は、m2処理前までロールバックすることにより、メッセージを正常な到着順序m1、m2、m3、m4、m5、m6に修正する。

【0087】

これにより、(3) 最上流の情報処理装置200Aでは、キュー201にメッセージm1、m2、m3、m4、m5、m6の順序で入り、マイクロサービス処理部101のメッセージm6のメッセージ処理をキャンセルする。また、中間の情報処理装置200Bでは、キュー201の修正対象メッセージm5を削除し、マイクロサービス処理部101はメッセージm4のメッセージ処理をキャンセルする。また、最下流の情報処理装置200Cでは、キュー201の修正対象メッセージm3を削除し、マイクロサービス処理部101はメッセージm2のメッセージ処理をキャンセルする。

【0088】

この際、メッセージの到着順序が正常でないと判断した情報処理装置200Aは、下流の全ての情報処理装置200B、200Cに対して上述した修正通知を送信する。

【0089】

これにより、上流の情報処理装置200Aでは、正常なメッセージの到着順序m1、m2、m3、m4、m5、m6でメッセージ処理することができる。また、中間および下流の情報処理装置200B、200Cでも、上流の情報処理装置200Aからメッセージm1、m2の順序でメッセージを受信し、正常にメッセージ処理できる。

【0090】

なお、この後に、たとえば、中間の情報処理装置200Bがメッセージを正常な到着順序で受信できない場合が生じれば、中間の情報処理装置200Bと下流の情報処理装置200C側は上記同様の修正処理を実行する。このように、複数の情報処理装置200のいずれの伝送路上でメッセージ遅延が生じた場合であっても、各情報処理装置200はメッセージの到着順序を修正してそれぞれ正常にメッセージ処理できるようになる。そして、図9に示すように、複数のマイクロサービスが連携し、情報処理装置200の段数が増え、影響を受けるメッセージの数が増えた場合でも同様の修正処理でメッセージの到着順序を正常に修正することができる。

【0091】

図10には、問題ケース4に対する修正処理例4を示す。この問題ケース4では、送信側2：受信側1のマイクロサービス間接続の例として、上流に複数のマイクロサービス(2つのマイクロサービス)が配置され、下流の一つのマイクロサービス(一つのマイクロサービス)に合流するシステム構成例を示す。

【0092】

図10には、上流側の2つのマイクロサービスのうち一方の情報処理装置200Aがメッセージm1～m4を送信し、他方の情報処理装置(不図示)がメッセージx1を送信する。そして、これらのメッセージm1～m4、x1が下流側の情報処理装置200Bで合流して受信する構成例を記載してある。

【0093】

そして、正常時の到着順序がメッセージm1、m2、x1、m3、m4であるとする。こ

10

20

30

40

50

ここで、（１）メッセージm１が遅延したため、到着順序がm２、x１、m３、m４、m１の順序に入れ替わった状態を示す。この際、下流の情報処理装置２００Ｂのキュー２０１にメッセージm２が入ったとする。また、上流の情報処理装置２００Ａでは、メッセージm２、m３のスナップショットを記録保持し、キュー２０１にはメッセージm４が入っている。

【００９４】

つぎに、（２）下流の情報処理装置２００Ｂでは、マイクロサービス処理部１０１がメッセージm２をメッセージ処理し、メッセージm２のスナップショットを記録保持する。また、上流の情報処理装置２００Ａでは、マイクロサービス処理部１０１がメッセージm４をメッセージ処理し、メッセージm２～m４のスナップショットを記録保持する。また、キュー２０１にメッセージm１が入る。

10

【００９５】

ここで、実施の形態では、情報処理装置２００Ａのメッセージ制御部１０２は、キュー２０１のメッセージm１のタイムスタンプと、スナップショットm２～m４の比較により、メッセージm１の到着遅れ（順序の逆転）を判断する。また、メッセージm２～m４が順序の修正対象であると判断する。そして、情報処理装置２００Ａのメッセージ制御部１０２は、m２処理前までロールバックすることにより、メッセージを正常な到着順序m１、m２、m３、m４に修正する。

【００９６】

これにより、（３）上流の情報処理装置２００Ａでは、キュー２０１にメッセージm１、m２、m３、m４の順序で入り、マイクロサービス処理部１０１はメッセージm４のメッセージ処理をキャンセルし、スナップショットm２～m４を削除する。

20

【００９７】

また、下流の情報処理装置２００Ｂでは、キュー２０１の修正対象メッセージm３を削除し、キュー２０１にはメッセージx１のみとなる。この際、マイクロサービス処理部１０１はメッセージm２のメッセージ処理をキャンセルし、メッセージ処理を停止（stop）する。

【００９８】

この際、メッセージの到着順序が正常でないと判断した情報処理装置２００Ａは、下流の情報処理装置２００Ｂに対して上述した修正通知を送信する。

30

【００９９】

つぎに、（４）上流の情報処理装置２００Ａでは、マイクロサービス処理部１０１がメッセージm１をメッセージ処理し、スナップショットm１を記録保持する。下流の情報処理装置２００Ｂでは、マイクロサービス処理部１０１はメッセージ処理が停止したままである。

【０１００】

この後、（５）上流の情報処理装置２００Ａでは、マイクロサービス処理部１０１がメッセージm２をメッセージ処理し、スナップショットm１、m２を記録保持する。キュー２０１には、メッセージm３、m４がある。下流の情報処理装置２００Ｂでは、メッセージ制御部１０２が受信メッセージを時間順でソートし、キュー２０１にはメッセージm１、x１の順に入る。この際、キュー２０１の先頭が修正後のメッセージ（m１）であれば、メッセージ制御部１０２は、マイクロサービス処理部１０１の停止を解除する。

40

【０１０１】

この後、（６）上流の情報処理装置２００では、正常化されたメッセージの到着順序m１、m２、m３、m４をメッセージ処理することができる。また、下流の情報処理装置２００Ｂでも、合流するメッセージm１、m２、x１、m３、m４の順序を有して正常にメッセージ処理できる。

【０１０２】

実施の形態では、上記の問題ケース１～４に基づき、スナップショットを保持するための最大の記録容量（サイズN_{max}）を下記式（１）に基づき設定する。

50

【 0 1 0 3 】

【 数 1 】

$$N_{max} = 1 + \sum_{i=1}^m (Q_i + 1) \quad \dots (1)$$

【 0 1 0 4 】

ただし、 m ：下流のマイクロサービス数（マイクロサービス処理部 1 0 1 の数）、 Q_i ＝下流 i 番目のマイクロサービスのキューサイズである。

10

【 0 1 0 5 】

図 1 1 は、実施の形態にかかる情報処理装置に必要なスナップショットのサイズを示す説明図である。3つのマイクロサービス処理部 1 0 1（情報処理装置 2 0 0 A～2 0 0 C）を有し、メッセージ到着順序が $m_1 \sim m_8$ 、 m_x である場合（メッセージ m_x はマイクロサービス処理前の状態）を例に説明する。

【 0 1 0 6 】

この場合、最上流の情報処理装置 2 0 0 A のマイクロサービス処理部 S 1（1 0 1）が修正をおこなうには、最大でメッセージ $m_1 \sim m_8$ が対象となる（スナップショットの数＝8）。中間の情報処理装置 2 0 0 B のマイクロサービス処理部 S 2（1 0 1）が修正をおこなうには、最大でメッセージ $m_1 \sim m_4$ が対象となる（スナップショットの数＝4）。最下流の情報処理装置 2 0 0 C のマイクロサービス処理部 S 3（1 0 1）が修正をおこなうには、最大でメッセージ m_1 が対象となる（スナップショットの数＝1）。

20

【 0 1 0 7 】

このように、実施の形態によれば、下流のマイクロサービス（マイクロサービス処理部 1 0 1）になるほどスナップショット記録数は少なくて済む。すなわち、実施の形態では、上流から下流までの全てのマイクロサービスのスナップショット記録数を同じ（最上流マイクロサービスで必要な記録数）にする必要はなく、システム全体で必要な記憶資源（履歴記録部 1 0 3）のサイズを小さくすることができる。

【 0 1 0 8 】

また、メッセージの到着遅れにより生じた間違ったマイクロサービス実行を正しいマイクロサービス実行にリカバリすることができ、かつ、履歴記録部 1 0 3 で必要なスナップショット記録数を必要最小限に抑えることができる。

30

【 0 1 0 9 】

また、メッセージ制御部 1 0 2 は、最初にメッセージの到着遅れを検知するまでは、スナップショットを記録せず、最初のメッセージ到着遅れ検出の後、下流のマイクロサービスも含めスナップショット記録を開始してもよい。この場合、メッセージ制御部 1 0 2 は、最初のメッセージの到着遅れによる誤ったマイクロサービス処理のリカバリ（修正処理）はおこなわないこととする。

【 0 1 1 0 】

また、メッセージ制御部 1 0 2（検査部 2 1 2）は、最初のメッセージ到着遅れを検知するまで、マイクロサービス処理部 1 0 1 が最後に処理したメッセージのタイムスタンプ t_a を保持する。そして、到着したメッセージのタイムスタンプ t_b と比較して、 t_b が t_a よりも古い場合、到着遅れと判定してもよい。

40

【 0 1 1 1 】

この場合、メッセージ到着遅れ検出後のリカバリは、2回目のメッセージ到着遅れからとなる。しかし、メッセージ制御部 1 0 2 は、全てのマイクロサービス処理部 1 0 1 でのスナップショットを記録する必要がなくなるため、スナップショット記録に必要な記憶資源（履歴記録部 1 0 3）のサイズを削減できる。

【 0 1 1 2 】

履歴記録部 1 0 3 は、マイクロサービスの初期状態のスナップショット $SS[0]$ と、処

50

理したメッセージだけを記録してもよい。スナップショット $SS[i]$ ($i = 1, \dots, n$) はメッセージの記録のみとなる。この場合、ロールバック制御部 213 によるロールバック処理は、マイクロサービスの状態を $SS[0]$ へ戻し、求めた $SS[r]$ までのメッセージを順番に実行させる。これにより、スナップショットのサイズを小さくすることができ、スナップショット記録に必要な記憶資源（履歴記録部 103）のサイズを削減できる。

【0113】

（従来技術について）

図12は、従来技術によるメッセージ通信の例を示す説明図である。図12(a)は、マイクロサービス1201と、マイクロサービス1202が1:1の場合のシステム1200の構成例、図12(b)は、 $n:1$ の場合のシステム1210の構成例を示す。

10

【0114】

図12(a)の1:1のシステム1200では、マイクロサービスA(1201)がメッセージa、bの順序で送信した後、経路上での通信の輻輳等によりメッセージaに遅延Xが生じたとする。この場合、たとえば、MQなどの従来の順序性保証機構により、マイクロサービスB(1202)では正常なメッセージa、bの順序で受信することができる。

【0115】

しかし、従来技術では、図12(b)の $n:1$ のシステム1210では、メッセージの順序性を保証することができない。上流に2つのマイクロサービスA1(1201a)、A2(1201b)が配置され、マイクロサービスA1(1201a)がメッセージaを送信し、マイクロサービスA2(1201b)がメッセージbを送信する。下流のマイクロサービスB(1202)は、上流の2つのマイクロサービスA1、A2(1201a、1201b)のメッセージa、bが合流する形で受信する。

20

【0116】

このような経路上で合流を含むシステム1210では、メッセージの順序性を保証できない。図12(b)の場合、下流側ではメッセージa、b、あるいはメッセージb、aの到着順序で受信するが、順序性保証機構を用いてもいずれの順序が正しいものであるか判断できない。

【0117】

通信の輻輳などでメッセージの到着順序が変わると、誤った処理結果がメッセージとして下流へ伝播し、下流全ての処理が誤った結果となる。従来技術では、古いメッセージが到着してはじめてメッセージの到着遅れを判断できるため、必ず誤った処理後のメッセージを送った後になり、下流へメッセージを送る前に止めることはできない。

30

【0118】

たとえば、システム1210においてマイクロサービスB(1202)が正しいメッセージa、bの順序で受信した場合には、この到着順序a、bで正しくメッセージ処理できる。しかし、メッセージaに遅延が生じて、マイクロサービスB(1202)がメッセージb、aの順序で受信した場合には、この到着順序b、aでメッセージ処理し下流に送信する。マイクロサービスB(1202)は、メッセージa、bが揃った後にタイムスタンプを比較する等してメッセージaの到着遅れを検出するが、この到着遅れの検出は、既に誤った処理結果を送出した後である。誤った処理結果はシステム1210の下流の全てのマイクロサービスに伝播してしまう。

40

【0119】

図13は、 $n:1$ のメッセージ通信で順序保証が必要な事象を示す説明図である。近年、IoT(Internet of Things)技術により実世界で発生した事象をクラウド上のメッセージ(イベント)に変換し、マイクロサービス間のメッセージ通信で処理するようになってきている。

【0120】

図13には、実世界で走行する車両の速度違反システムの例を示す。実世界では車両に設けられたセンサ等により車両の現在位置(一般道または高速道路)と、車両の速度を検出

50

する。クラウド側のシステム1300は、マイクロサービスA1(1301a)が車両の現在位置のメッセージ処理をおこない、マイクロサービスA2(1301b)が車両の速度検出のメッセージ処理をおこなう。そして、マイクロサービスB(1302)は、マイクロサービスA1、A2(1301a、1301b)のメッセージを合流して受信し、スピード違反のメッセージ処理をおこない、マイクロサービスC(1303)は免許証の減点のメッセージ処理をおこなう。

【0121】

上記マイクロサービスのシステム1300において、実世界での車両の現在位置と速度は、正しくはメッセージa、bの順序であったとする。メッセージaは現在位置が高速道路であり、メッセージbが速度100km/hであるとする。

【0122】

しかし、メッセージaに遅延Xが生じた場合、マイクロサービスB(1302)では誤った到着順序b、aでメッセージ処理してしまう。この場合、実世界の車両は高速道路上で100km/hの速度で走行しているところ、マイクロサービスB(1302)では車両が高速道路ではない現在位置(一般道)上で速度100km/hで走行するため、スピード違反であるとメッセージ処理してしまう。さらに、下流のマイクロサービスC(1303)では、スピード違反のメッセージ処理に基づき所定の減点のメッセージ処理をしてしまう。

【0123】

このように、クラウド上のシステム1300では、実世界の事象発生に伴うメッセージを正しい順序で伝播させないと誤った処理結果を生み出し、一つのマイクロサービスで生じた誤りをさらに下流のマイクロサービスに伝播させてしまう。

【0124】

図14A、図14Bは、従来技術による複数のマイクロサービスを連携させたときの問題を示す説明図である。これら図14A、図14Bには、ある部屋の省エネ効果測定を5つのマイクロサービスで連携して処理するシステムを示す。図14Aは、マイクロサービス間でのメッセージ順序が正しい場合、図14Bは、マイクロサービス間でのメッセージ順序が入れ替わった場合を示す。

【0125】

図14Aの(a)は、システム1400の構成例であり、図14Aの(a)に示す実世界に配置した入退室検知器Aは部屋へ入室した人体の数を入退室カード等によりユーザ別(個別)に検出し、電力計測器Bは部屋で消費する電力を検出する。システム1400の5つのマイクロサービスは、在室者カウンタ(マイクロサービス)1401が入退室検知器Aの検知情報に基づき、部屋への人体の入退室をカウント(加減算)し、現在部屋にいる人体の数を判断するメッセージ処理をおこなう。消費電力計測(マイクロサービス)1402は、電力計測器Bの計測の情報に基づき、部屋の消費電力を計測するメッセージ処理をおこなう。

【0126】

効率計算(マイクロサービス)1403は、在室者カウンタ1401と、消費電力計測1402のマイクロサービスの合流したメッセージを受信し、現在部屋にいる人体の数あたりの電力消費の効率(W/人)を求めるメッセージ処理をおこなう。省エネ評価(マイクロサービス)1404は、効率計算の結果に基づく省エネ状態を点数で評価するメッセージ処理をおこなう。月間効果測定(マイクロサービス)1405は、部屋の月間の省エネ効果の評価(たとえば、点数の推移、月間の評価の点数の平均、前回評価の点数に対する差分等)のメッセージ処理をおこなう。

【0127】

図14Aの(b)には、メッセージの到着順序が正常時のシステム1400の各マイクロサービスの動作例を示すシーケンス図を示す。はじめに、入退室検知器Aが部屋から一人の退室を検知し、このときの在室者カウンタ(マイクロサービス)1401が在室者一人と判断したとする。そして、電力計測器Bが200Wの消費電力を計測すると、効率計算

10

20

30

40

50

(マイクロサービス) 1403は200W/人の効率計算をおこなう。そして、省エネ評価(マイクロサービス) 1404は在室者の省エネ評価を60点と評価する。月間効果測定(マイクロサービス) 1405は前回の省エネ評価時 t_1 ～今回の省エネ評価時 t_2 の点数推移を求める。たとえば前回の点数が80点だった場合、 t_1 ～ t_2 が80点として出力する。

【0128】

この後、入退室検知器Aが部屋へ一人の入室を検知し、このときの在室者カウンタ(マイクロサービス) 1401が在室者2人と判断したとすると、効率計算(マイクロサービス) 1403は100W/人の効率計算をおこなう。そして、省エネ評価(マイクロサービス) 1404は在室者の省エネ評価を100点と評価する。月間効果測定(マイクロサービス) 1405は t_2 ～ t_3 が60点である情報(メッセージ)を出力する。

10

【0129】

つぎに、図14Bに示すマイクロサービス間でのメッセージの到着順序が入れ替わった場合について説明する。図14Bの(a)に示すように、システム1400の在室者カウンタ(マイクロサービス) 1401のメッセージに遅延Xが発生したとする。

【0130】

図14Bの(b)に示すように、遅延Xの発生が一人入室時に生じた場合、在室者カウンタ(マイクロサービス) 1401から出力されるメッセージは、効率計算(マイクロサービス) 1403に本来の到達時点(時期 t_0)で受信される。しかし、遅延Xにより、在室者カウンタ(マイクロサービス) 1401から出力されるメッセージが実際には時期 t_x で効率計算(マイクロサービス) 1403に到達(受信)したとする。

20

【0131】

この場合、遅延Xによる期間Tの間、効率計算(マイクロサービス) 1403は、一人の入室にかかるメッセージを受信しない状態のままメッセージ処理をおこなう。この場合、本来入室は二人であるところ、一人しか入室していない状態(合計一人)だけで電力計測器Bが検出した消費電力800Wに基づく効率計算(800W/一人)をおこなう。この結果、下流の省エネ評価(マイクロサービス) 1404は10点と評価し、月間効果測定(マイクロサービス) 1405も t_2 ～ t_4 が60点と判断してしまう。図中×印は、本来必要なメッセージ処理であるが未処理となったメッセージ処理結果、△印は誤った処理によるメッセージ処理結果である。

30

【0132】

遅延Xのメッセージによる期間T経過後、システム1400の消費電力計測(マイクロサービス)は、遅延したメッセージが時期 t_x で実際に到着した時点ではじめてこのメッセージの到着遅れを検出する。このように、従来技術では、メッセージに遅延が生じた場合、遅延した期間Tまでの間、メッセージの遅延の有無を判断できずにメッセージ処理をおこなってしまい、下流の全てのマイクロサービスも同様に誤ったメッセージ処理をおこなう問題が生じた。

【0133】

これに対し、実施の形態によれば、各マイクロサービスにマイクロサービス間のメッセージの伝播状態と、メッセージ処理時のタイムスタンプによりメッセージの到着遅れを検知できる。一つのマイクロサービスのメッセージ遅延はキュー内でのタイムスタンプの比較だけで簡単に検出でき、正しいメッセージ順序に修正できる。

40

【0134】

また、メッセージの遅延が複数のマイクロサービスに影響を与える場合、たとえば、遅延したメッセージのメッセージ処理を実行せずに下流に処理後のメッセージを出した場合にも対応できる。この場合、メッセージ処理後のスナップショットを参照して、ロールバック制御することで、正しいメッセージ順序に修正でき、再度メッセージ処理できるようになる。具体的には、下流のマイクロサービスでは、到着遅れメッセージより先行したメッセージをメッセージ処理した以前の状態のスナップショットを参照して状態を戻し、到着遅れのメッセージから正しいメッセージ順序でメッセージ処理を再実行することができる。

50

【0135】

また、上流の複数のマイクロサービスのメッセージが合流して下流の一つのマイクロサービスで受信するシステム構成においても、修正通知によりサービス処理部でのメッセージ処理を停止する。このように上流で並列にメッセージ処理する場合において、たとえば一方のメッセージに到着遅れが生じて、自身のマイクロサービスでは正しいメッセージ到着順序に修正できる。加えて、下流のマイクロサービスでは、メッセージ処理を停止することで誤ったメッセージ処理結果をさらに下流に伝播させることを防ぐことができる。また、キューの先頭のメッセージが正しく修正されたメッセージになれば停止を解除して、メッセージ処理を再開できる。

【0136】

また、サービス処理部の処理時のスナップショットは、キューサイズから求めた所定数分だけ保存すればよい。この際、保持するスナップショットは、下流の情報処理装置のサービス処理部の数と、キューサイズに基づく記録容量とすることができ、下流のマイクロサービスに至るにつれて記録容量を小さくできる。これにより、システム全体でのスナップショットを保持するための記録容量を削減できるようになる。

【0137】

従来スナップショットを保持する記録容量は、運用者が勘により経験上十分と考える世代数分にするため、余裕を持った世代数となり、必要以上のスナップショットを保存していた。これに対し、実施の形態では、スナップショットの記録容量を削減でき、記憶資源にかかるコストを低コスト化できるようになる。

【0138】

そして、実施の形態によれば、上流のマイクロサービスで到着遅れを検出して下流のマイクロサービスに修正通知を送信するため、簡単な構成で正しいメッセージ順序に修正できる。このため、メッセージの順序保証を伝送路等でおこなう必要がなく、高価な専用システムを用いる必要がない。

【0139】

そして、実施の形態によれば、所定のサービスを提供する複数のマイクロサービスのいずれかでメッセージ到着遅れが生じて、簡単な処理と低コストな構成で全てのマイクロサービスがそれぞれ正しいメッセージ到着順序でメッセージ処理できるようになる。

【0140】

なお、本発明の実施の形態で説明した情報処理プログラムは、あらかじめ用意されたプログラムをサーバ等のプロセッサに実行させることにより実現することができる。本制御方法は、ハードディスク、フレキシブルディスク、CD-ROM (Compact Disc-Read Only Memory)、DVD (Digital Versatile Disk)、フラッシュメモリ等のコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また本制御方法は、インターネット等のネットワークを介して配布してもよい。

【0141】

上述した実施の形態に関し、さらに以下の付記を開示する。

【0142】

(付記1) 複数のマイクロサービス間でメッセージ通信し、それぞれが前記マイクロサービスをおこなう情報処理装置において、
受信した複数の前記メッセージのタイムスタンプを比較し、遅延して到着した前記メッセージを判断し、前記メッセージを正しい到着順序で再度メッセージ処理し、前記メッセージを受信する下流の情報処理装置に対し、正しいメッセージの到着順序で再度メッセージ処理させる修正通知を送信する、
メッセージ制御部を有することを特徴とする情報処理装置。

【0143】

(付記2) 受信したメッセージを到着順序で格納するキューと、
前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサ

10

20

30

40

50

ービス処理部と、を有し、
前記メッセージ制御部は、
前記サービス処理部のメッセージ処理毎の処理状態をスナップショットとして処理順に保持する履歴記録部と、
前記キューに格納された前記メッセージと、前記履歴記録部に保持された前記スナップショットに基づき、前記受信したメッセージの到着順序が正常か否かを検査し、検査結果を前記修正通知として送信する検査部と、
前記修正通知の受信時に、前記キューおよび前記サービス処理部に対し、正しいメッセージの到着順序で処理をおこなわせるロールバック制御をおこなうロールバック制御部と、
を有する、
ことを特徴とする付記 1 に記載の情報処理装置。

10

【0144】

(付記 3) 前記検査部は、
下流の情報処理装置に対する前記修正通知として、修正対象の前記メッセージの識別子と、到着順序を正しく修正した後のメッセージ列の先頭のメッセージの識別子の情報を送信することを特徴とする付記 2 に記載の情報処理装置。

【0145】

(付記 4) 前記ロールバック制御部は、
前記修正通知の受信時に、当該修正通知に基づき、前記キューに対し、正しく修正した後のメッセージ列を格納し、前記サービス処理部に対し、遅延して到着した前記メッセージ以前のメッセージ処理状態にロールバックさせることを特徴とする付記 2 または 3 に記載の情報処理装置。

20

【0146】

(付記 5) 前記ロールバック制御部は、
前記修正通知の受信時に、当該修正通知に基づき、前記サービス処理部に対し、到着順序が正しくない前記メッセージのメッセージ処理をキャンセルさせることを特徴とする付記 4 に記載の情報処理装置。

【0147】

(付記 6) 前記ロールバック制御部は、
前記修正通知の受信時に、前記キューに格納された修正対象の前記メッセージを削除することを特徴とする付記 4 または 5 に記載の情報処理装置。

30

【0148】

(付記 7) 上流の複数の情報処理装置から合流して前記メッセージを受信するシステム構成の場合、

前記ロールバック制御部は、

前記修正通知の受信時に、前記サービス処理部によりメッセージ処理を停止し、前記キューの先頭のメッセージが正しく修正されたメッセージであれば前記停止を解除することを特徴とする付記 2 または 3 に記載の情報処理装置。

【0149】

(付記 8) 前記履歴記録部は、

下流の情報処理装置の前記サービス処理部の数と、キューサイズに基づく記録容量を有することを特徴とする付記 2 ～ 7 のいずれか一つに記載の情報処理装置。

40

【0150】

(付記 9) 前記履歴記録部は、

前記スナップショットとして、前記サービス処理部のメッセージ処理毎のサービス状態の履歴を記録保持することを特徴とする付記 2 ～ 8 のいずれか一つに記載の情報処理装置。

【0151】

(付記 10) 複数のマイクロサービス間でメッセージ通信し、それぞれが前記マイクロサービスをおこなう複数の情報処理装置を含む情報処理システムにおいて、
前記情報処理装置は、それぞれ、

50

受信したメッセージを到着順序で格納するキューと、
前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、
前記メッセージのタイムスタンプを比較し、遅延して到着した前記メッセージを判断し、
前記メッセージを正しい到着順序で再度メッセージ処理し、前記メッセージを受信する下流の情報処理装置に対し、正しいメッセージの到着順序で再度メッセージ処理させる修正通知を送信するメッセージ制御部と、
を有することを特徴とする情報処理システム。

【0152】

（付記11）複数のマイクロサービス間でメッセージ通信し、それぞれが前記マイクロサービスをおこなう複数の情報処置装置と、メッセージ制御装置とを含む情報処理システムにおいて、

前記情報処理装置は、それぞれ、

受信したメッセージを到着順序で格納するキューと、

前記キューへの格納順に前記メッセージを読み出し、所定のメッセージ処理をおこなうサービス処理部と、を有し、

前記メッセージ制御装置は、

前記情報処理装置のメッセージのタイムスタンプを比較し、遅延して到着した前記メッセージを判断し、前記メッセージを正しい到着順序で再度メッセージ処理し、前記メッセージを受信する下流の情報処理装置に対し、正しいメッセージの到着順序で再度メッセージ処理させる修正通知を送信する、
ことを特徴とする情報処理システム。

【0153】

（付記12）複数のマイクロサービス間でメッセージ通信し、それぞれが前記マイクロサービスをおこなう情報処置装置のコンピュータに、

受信した複数の前記メッセージのタイムスタンプを比較し、遅延して到着した前記メッセージを判断し、

前記メッセージを正しい到着順序で再度メッセージ処理し、

前記メッセージを受信する下流の情報処理装置に対し、正しいメッセージの到着順序で再度メッセージ処理させる修正通知を送信する、

処理を実行させることを特徴とする情報処理プログラム。

【0154】

（付記13）前記修正通知の受信時に、当該修正通知に基づき、正しく修正した後のメッセージ列をキューに格納し、

前記マイクロサービスの処理をおこなうサービス処理部に対し、遅延して到着した前記メッセージ以前のメッセージ処理状態にロールバックさせる、

処理を実行させることを特徴とする付記12に記載の情報処理プログラム。

【符号の説明】

【0155】

100 情報処理システム

101 マイクロサービス処理部

102 メッセージ制御部

103 履歴記録部

200 (200A～200C) 情報処理装置

201 キュー

212 検査部

213 ロールバック制御部

301 CPU

302 メモリ

303 ネットワークインタフェース

10

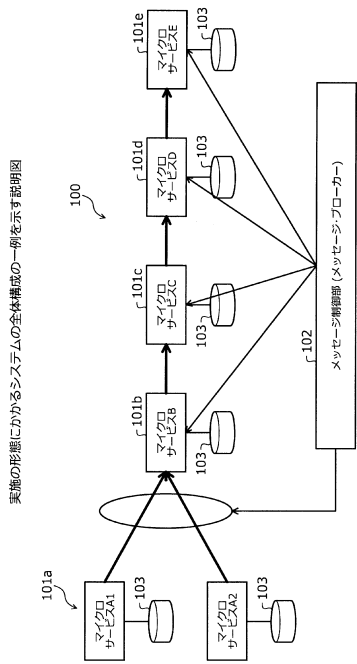
20

30

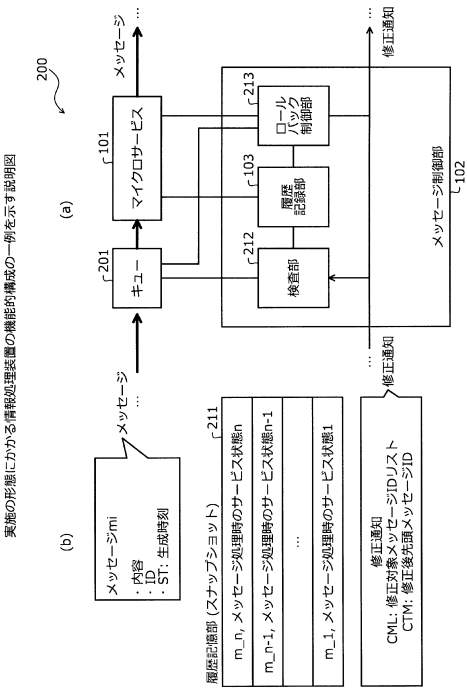
40

50

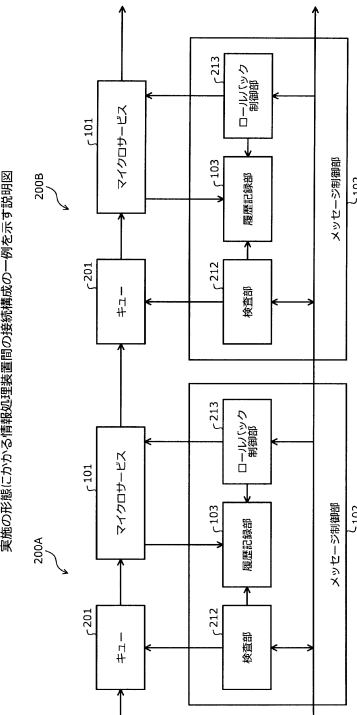
3 0 5 記録媒体
3 3 0 ネットワーク
【図面】
【図 1】



【図 2 A】

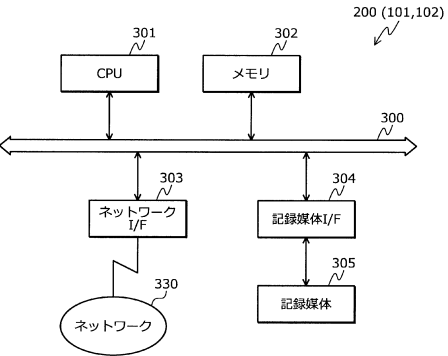


【図 2 B】



【図 3】

実施の形態にかかる情報処理装置のハードウェア構成の一例を示すブロック図



10

20

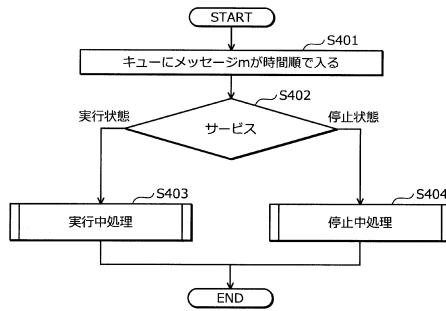
30

40

50

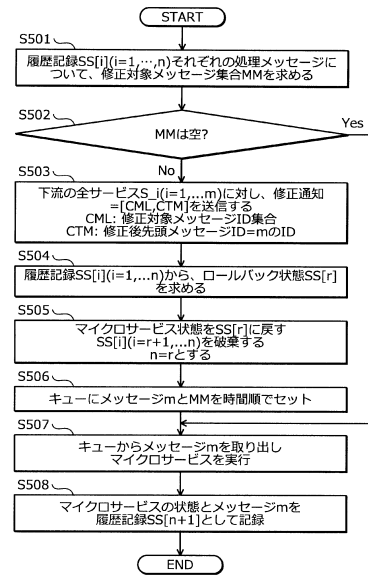
【図 4】

実施の形態にかかる情報処理装置がおこなう全体の処理の手順の一例を示すフローチャート



【図 5】

実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャート
(その1)

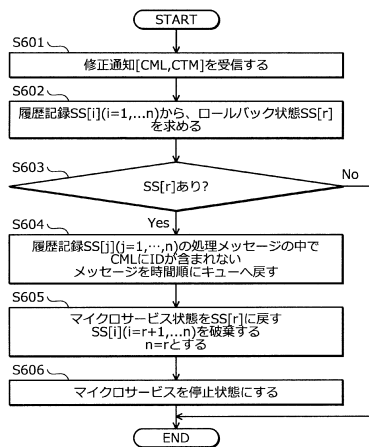


10

20

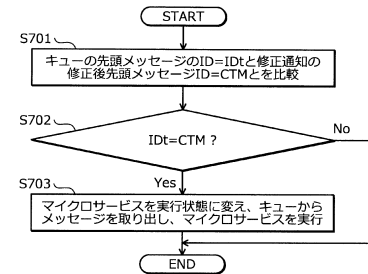
【図 6】

実施の形態にかかる情報処理装置の実行中処理の手順の一例を示すフローチャート
(その2)



【図 7】

実施の形態にかかる情報処理装置の停止中処理の手順の一例を示すフローチャート



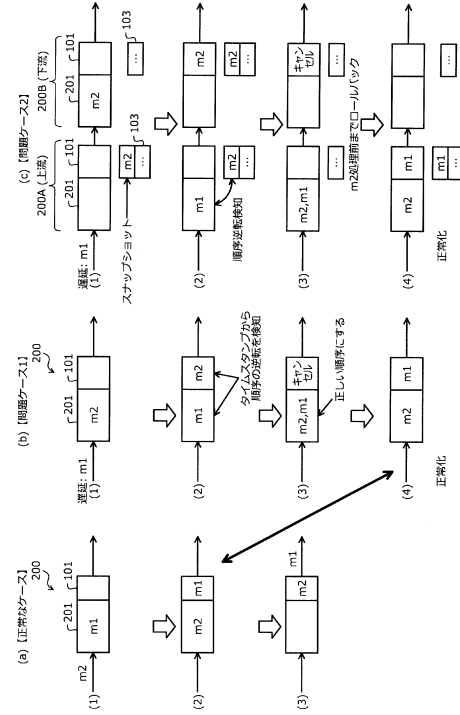
30

40

50

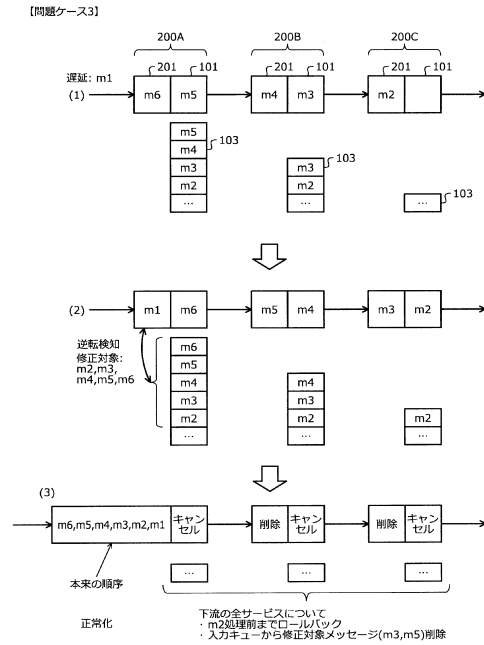
【図 8】

実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その 1）



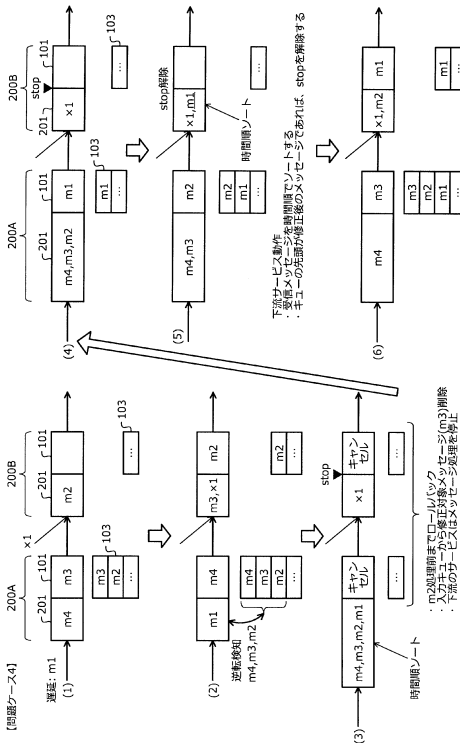
【図 9】

実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その 2）



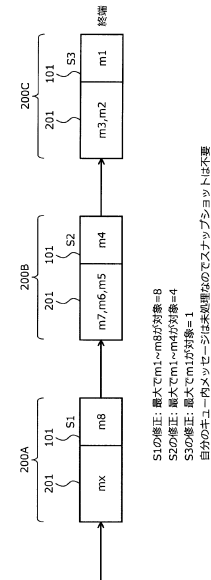
【図 10】

実施の形態にかかる情報処理装置のメッセージ処理の一例を示す説明図（その 3）



【図 11】

実施の形態にかかる情報処理装置に必要なスナップショットのサイズを示す説明図



10

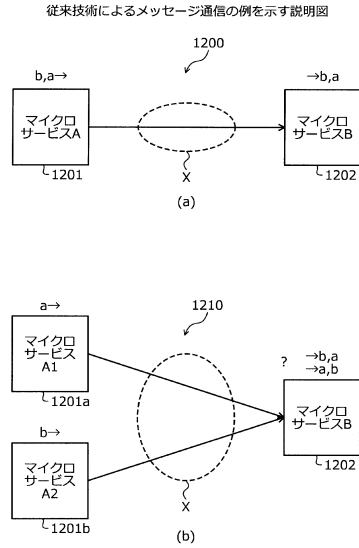
20

30

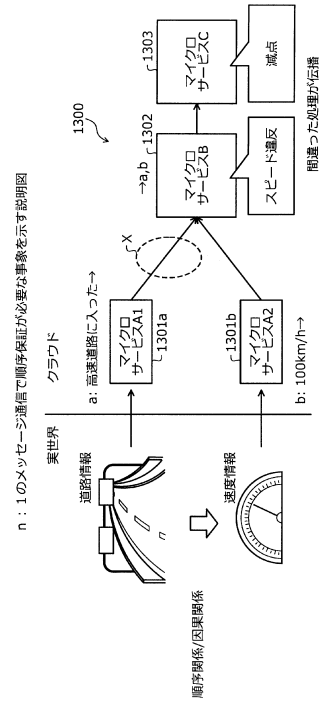
40

50

【図 1 2】



【図 1 3】

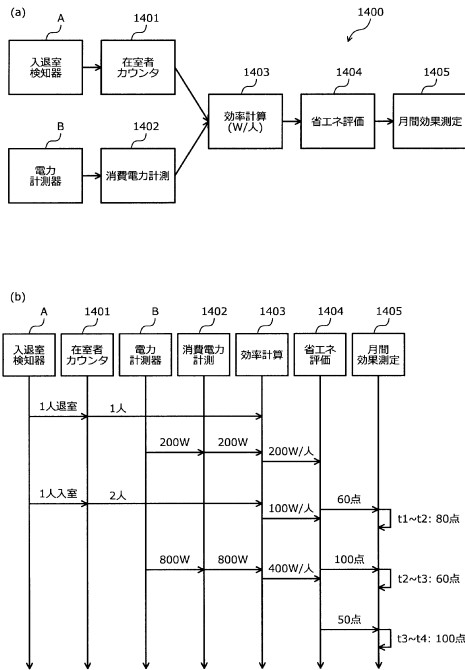


10

20

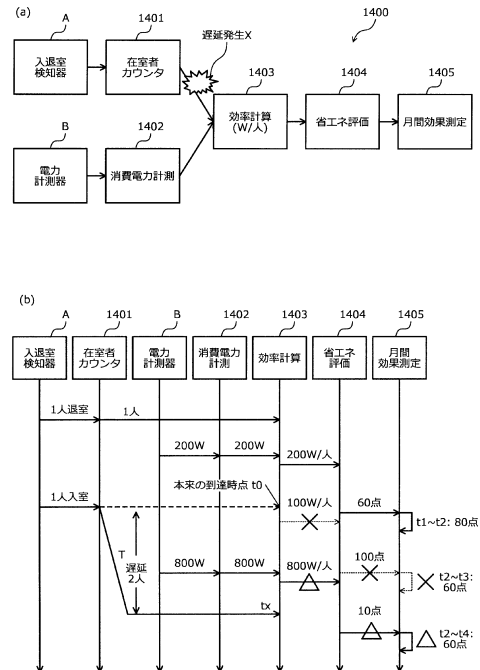
【図 1 4 A】

従来技術による複数のマイクロサービスを連携させたときの問題を示す説明図 (その 1)



【図 1 4 B】

従来技術による複数のマイクロサービスを連携させたときの問題を示す説明図 (その 2)



30

40

50

フロントページの続き

審査官 三坂 敏夫

(56)参考文献 特開平 06-028199 (JP, A)

特開平 11-259386 (JP, A)

特開平 10-283276 (JP, A)

山口 佳紀 他, 実時間履歴情報を用いた並列離散事象シミュレーション, 電気学会研究会資料, 社団法人電気学会, 1997年01月31日, 第97-1/11巻, 第41頁-第46頁

(58)調査した分野 (Int.Cl., DB名)

G 0 6 F 9 / 4 5 5 - 9 / 5 4