

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-9852

(P2008-9852A)

(43) 公開日 平成20年1月17日(2008.1.17)

(51) Int. Cl.

G06F 9/50 (2006.01)

F I

G06F 9/46 465 D

テーマコード (参考)

審査請求 未請求 請求項の数 9 O L (全 13 頁)

(21) 出願番号 特願2006-181428 (P2006-181428)
 (22) 出願日 平成18年6月30日 (2006.6.30)

(71) 出願人 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100123788
 弁理士 宮崎 昭夫
 (74) 代理人 100106138
 弁理士 石橋 政幸
 (74) 代理人 100127454
 弁理士 緒方 雅昭
 (72) 発明者 岡田 真明
 東京都港区芝五丁目7番1号 日本電気株式会社内

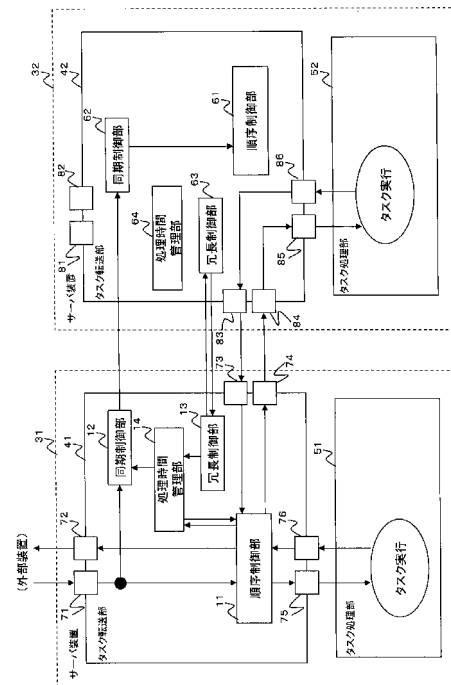
(54) 【発明の名称】 負荷分散制御システム、方法、およびサーバ装置

(57) 【要約】

【課題】 実処理だけでなく伝送路の負荷をも考慮してネットワークに接続された負荷を分散することができ、応答の即時性を向上させることができる負荷分散制御システムを提供する。

【解決手段】 サーバ装置31のタスク転送部41がタスク処理部51の実処理の負荷および他サーバ装置32の実処理の負荷に加えて、他サーバ装置32との間の伝送路の負荷を監視し、それら実処理の負荷と伝送路の負荷に基づいてコマンド・データの割り当て先を決定する。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

負荷を分散して処理を実行する負荷分散制御システムであって、
伝送路を介して相互に接続され、それぞれがコマンドの実処理を実行する複数のサーバ装置を有し、前記複数のサーバ装置の少なくとも 1 つが、
割り当てられたコマンドの実処理を実行する処理部と、
前記処理部の実処理の負荷および他サーバ装置の実処理の負荷と、前記他サーバ装置との間の前記伝送路の負荷を監視しており、処理要求のコマンドを受けると、前記処理部の実処理の負荷と前記他サーバ装置の実処理の負荷と前記伝送路の負荷とに基づいて前記コマンドの割り当て先を決め、該コマンドを該割り当て先に転送する転送部と、を有する負荷分散制御システム。 10

【請求項 2】

前記転送部は、前記コマンドを転送した時刻と前記コマンドの処理結果を受けた時刻とから前記実処理の所要時間を算出し、該実処理の所要時間で前記実処理の負荷の状態を判断する請求項 1 記載の負荷分散制御システム。

【請求項 3】

前記転送部は、前記他のサーバ装置にヘルスチェックデータを送信し、該他のサーバ装置から該ヘルスチェックデータに対する応答を受信し、該ヘルスチェックデータを送信した時刻と前記応答を受信した時刻とからヘルスチェックの所要時間を算出し、該ヘルスチェックの所要時間で前記伝送路の負荷の状態を判断する請求項 1 または 2 記載の負荷分散制御システム。 20

【請求項 4】

前記転送部は、前記他サーバ装置にヘルスチェックデータを送信した後、所定の時間内に該他サーバ装置から該ヘルスチェックデータに対する応答を受信できなければ、該他サーバ装置を前記コマンドの割り当て先の候補から除外する、請求項 3 に記載の負荷分散制御システム。

【請求項 5】

負荷を分散して処理を実行する負荷分散制御方法であって、
伝送路を介して相互に接続され、それぞれがコマンドの実処理を実行する複数のサーバ装置を有し、前記複数のサーバ装置の少なくとも 1 つが、
割り当てられたコマンドの実処理を実行し、
前記実処理の負荷および他サーバ装置の実処理の負荷と、前記他サーバ装置との間の前記伝送路の負荷を監視し、処理要求のコマンドを受けると、前記実処理の負荷と前記他サーバ装置の実処理の負荷と前記伝送路の負荷とに基づいて前記コマンドの割り当て先を決め、該コマンドを該割り当て先に転送する負荷分散制御方法。 30

【請求項 6】

前記コマンドを転送した時刻と前記コマンドの処理結果を受けた時刻とから前記実処理の所要時間を算出し、該実処理の所要時間で前記実処理の負荷の状態を判断する請求項 5 記載の負荷分散制御方法。

【請求項 7】

前記他のサーバ装置にヘルスチェックデータを送信し、該他のサーバ装置から該ヘルスチェックデータに対する応答を受信し、該ヘルスチェックデータを送信した時刻と前記応答を受信した時刻とからヘルスチェックの所要時間を算出し、該ヘルスチェックの所要時間で前記伝送路の負荷の状態を判断する請求項 5 または 6 に記載の負荷分散制御方法。 40

【請求項 8】

前記他サーバ装置にヘルスチェックデータを送信した後、所定の時間内に該他サーバ装置から該ヘルスチェックデータに対する応答を受信できなければ、該他サーバ装置を前記コマンドの割り当て先の候補から除外する、請求項 7 に記載の負荷分散制御方法。

【請求項 9】

複数のサーバ装置が伝送路を介して接続された負荷分散制御システムを構成するサーバ 50

装置であって、

割り当てられたコマンドの実処理を実行する処理部と、

前記処理部の実処理の負荷および他のサーバ装置の実処理の負荷と、前記他サーバ装置との間の前記伝送路の負荷を監視しており、処理要求のコマンドを受けると、前記処理部の実処理の負荷と前記他サーバ装置の実処理の負荷と前記伝送路の負荷とに基づいて前記コマンドの割り当てを決め、該コマンドを該割り当て先に転送する転送部と、を有するサーバ装置。

【発明の詳細な説明】

【技術分野】

【0001】

10

本発明は、負荷分散の制御に関し、特に冗長構成のサーバ群の負荷分散の制御に関する。

【背景技術】

【0002】

従来、冗長構成のサーバ群の負荷分散において、各サーバ装置は自己の負荷状況を測定して測定結果を相互に通知し、サーバ装置間でその測定結果を比較している。そして、比較した結果から負荷の最も軽いサーバ装置を選択して処理が割り当てられる。

【0003】

特許文献1は、CPU負荷の軽いサーバ装置を動的に選択する負荷分散の技術を開示している。特許文献1では、クライアントは、ある特定のサーバ装置（以下「要求サーバ先装置」という）に対してデータベースへのアクセスを要求すると同時に他の全てのサーバ装置にCPU負荷情報を要求先サーバ装置に通知するように要求する。データベースへのアクセスの要求を受けた要求先サーバ装置は、他のサーバ装置から現在のCPU負荷情報を受信して、他のサーバ装置のCPUと自己のCPU負荷とを比較する。自己のCPU負荷が最も軽ければ、要求先サーバ装置自身がデータベースにアクセスする。他のサーバ装置のCPU負荷の方が軽ければ、要求先サーバ装置から最もCPU負荷の軽いサーバ装置に制御を移し、移動先サーバ装置がデータベースにアクセスする。

20

【特許文献1】特開平11-65912号公報

【発明の開示】

【発明が解決しようとする課題】

30

【0004】

特許文献1に開示されている技術では、クライアントが要求先サーバ装置にデータベースのアクセスを要求し、他のサーバ装置にCPU負荷情報の通知を要求すると、他の各サーバ装置はCPU負荷情報を生成し、要求先サーバ装置に通知する。要求先サーバ装置は、自身のCPU負荷と他のサーバ装置から通知されたCPU負荷情報に基づいて、データベースにアクセスするサーバ装置を決定し、決定されたサーバ装置がデータベースへのアクセスを開始していた。

【0005】

負荷分散を行う目的の1つにクライアントからのアクセスに対する応答の即時性の確保があるが、クライアントがアクセスを要求してから応答が変わるまでの時間を決める要因はCPU負荷だけではない。移動先サーバ装置のCPU負荷が軽くても、クライアントと移動先サーバ装置の間の伝送路が輻輳していれば、クライアントに応答が返るまでの時間が長くなることがあった。

40

【0006】

本発明の目的は、実処理だけでなく伝送路の負荷をも考慮してネットワークに接続された負荷を分散することができ、応答の即時性を向上させることができる負荷分散制御システムを提供することにある。

【課題を解決するための手段】

【0007】

上記目的を達成するために、本発明の負荷分散制御システムは、

50

負荷を分散して処理を実行する負荷分散制御システムであって、
伝送路を介して相互に接続され、それぞれがコマンドの実処理を実行する複数のサーバ装置を有し、前記複数のサーバ装置の少なくとも１つが、
割り当てられたコマンドの実処理を実行する処理部と、
前記処理部の実処理の負荷および他サーバ装置の実処理の負荷と、前記他サーバ装置との間の前記伝送路の負荷を監視しており、処理要求のコマンドを受けると、前記処理部の実処理の負荷と前記他サーバ装置の実処理の負荷と前記伝送路の負荷とに基づいて前記コマンドの割り当て先を決め、該コマンドを該割り当て先に転送する転送部と、を有する。

【０００８】

本発明によれば、転送部が処理部の実処理の負荷および他サーバ装置の実処理の負荷に加えて、他サーバ装置との間の伝送路の負荷を監視し、それら実処理の負荷と伝送路の負荷に基づいてコマンド・データの割り当て先を決定する。そのため、実処理だけでなく伝送路の負荷をも考慮してネットワークに接続された負荷を分散することができ、応答の即時性を向上させることができる。

【発明の効果】

【０００９】

本発明によれば、実処理だけでなく伝送路の負荷をも考慮してネットワークに接続された負荷を分散することができ、応答の即時性を向上させることができる。

【発明を実施するための最良の形態】

【００１０】

本発明を実施するための形態について図面を参照して詳細に説明する。

【００１１】

図１は、本実施形態における負荷分散制御システムの構成を示すブロック図である。本実施形態の負荷分散制御システムは、ノード装置１内にネットワーク装置２１～２ｎとサーバ装置３１，３２を有している。なお、ここでは一例としてサーバ装置が２つの構成を示しているが、３つ以上のサーバ装置が相互に接続された構成であってもよい。

【００１２】

ネットワーク装置２１～２ｎは、複数のサーバ装置３１，３２を相互に接続し、サーバ装置３１，３２間の通信を可能にする。

【００１３】

各サーバ装置３１，３２は、ネットワーク装置２１～２ｎを介して相互に接続され、冗長構成となっている。

【００１４】

図２は、本実施形態のサーバ装置３１，３２の冗長構成を示すブロック図である。サーバ装置３１は、タスク転送部４１およびタスク処理部５１を有している。サーバ装置３２は、タスク転送部４２およびタスク処理部５２を有している。サーバ装置３１において、タスク転送部４１は、順序制御部１１、同期制御部１２、冗長制御部１３、処理時間管理部１４、およびインタフェース７１，７２，７３，７４，７５，７６を有している。サーバ装置３２において、タスク転送部４２は、順序制御部６１、同期制御部６２、冗長制御部６３、処理時間管理部６４、およびインタフェース８１，８２，８３，８４，８５，８６を有している。

【００１５】

サーバ装置３１は、外部装置（不図示）からコマンド・データを受信して、自身で実処理を行うか、あるいは伝送路を介して接続されているサーバ装置３２にコマンド・データを転送する。

【００１６】

サーバ装置３２は、サーバ装置３１から転送されたコマンド・データを受信し、受信したコマンド・データの実処理を行い、処理結果の応答をサーバ装置３１に返す。

【００１７】

サーバ装置３１とサーバ装置３２は、同一の構成を有しており、外部装置からのコマン

10

20

30

40

50

ド・データをサーバ装置 3 2 で受信することも可能である。例えば、サーバ装置 3 1 が故障したとき、サーバ装置 3 2 はサーバ装置 3 1 と同様の処理を行うことができる。

【0018】

順序制御部 1 1 は、外部装置から受信したコマンド・データに対して処理の順序性を保証する制御を行う。順序制御部 1 1 は、外部装置から外部インタフェース 7 1 を経由して受信したコマンド・データをキューイングおよびバッファリングする。

【0019】

そして、順序制御部 1 1 は、処理時間管理部 1 4 から他のサーバ装置 3 2 が送信先として指定されていなければ、バッファリングしたコマンド・データを、インタフェース 7 5 を経由してタスク処理部 5 1 に送る。コマンド・データをタスク処理部 5 1 に送信した後、順序制御部 1 1 は処理開始通知情報を処理時間管理部 1 4 に送る。処理開始通知情報は、タスク処理部 5 1, 5 2 に実処理を要求したことを示す情報である。

10

【0020】

また、順序制御部 1 1 は、タスク処理部 5 1 からインタフェース 7 6 を経由して処理結果を受け取る。タスク処理部 5 1 から処理結果を受けると、順序制御部 1 1 は、処理終了通知情報を処理時間管理部 1 4 に送るとともに、処理結果をインタフェース 7 2 を介して外部装置に送る。処理終了通知情報は、タスク処理部 5 1, 5 2 が実処理を終了したことを示す情報である。

【0021】

順序制御部 1 1 は、処理時間管理部 1 4 から他のサーバ装置 3 2 が送信先として指定されていれば、インタフェース 7 4 を経由してサーバ装置 3 2 のタスク転送部 4 2 にコマンド・データを転送する。そして、順序制御部 1 1 は、サーバ装置 3 2 のタスク転送部 4 2 からインタフェース 7 3 を経由してコマンド・データの処理結果を受けると、サーバ装置 3 2 の処理終了通知情報を処理時間管理部 1 4 に送るとともに、処理結果をインタフェース 7 2 を介して外部装置に返送する。

20

【0022】

その際、サーバ装置 3 1 とサーバ装置 3 2 はそれぞれ独立に処理を実行するので、順序制御部 1 1 の入力で見ると、サーバ装置 3 1 の処理結果とサーバ装置 3 2 の処理結果の順序性は保証されていない。そこで、順序制御部 1 1 は、サーバ装置 3 1 の処理結果とサーバ装置 3 2 の処理結果の順序を正常にして外部装置に返送する。

30

【0023】

同期制御部 1 2 は、冗長構成のサーバ装置 3 1 とサーバ装置 3 2 の間の処理の同期をとる。同期制御部 1 2 は、インタフェース 7 1 からのコマンド・データを受信する。また、同期制御部 1 2 は、処理時間管理部 1 4 から処理開始情報と各サーバ装置 3 1, 3 2 の処理終了通知情報を受信する。同期制御部 1 2 は、インタフェース 7 1 からのコマンド・データと処理時間管理部 1 4 からの処理開始通知情報とを他のサーバ装置 3 2 の同期制御部 6 2 に送信する。サーバ装置 3 2 の同期制御部 6 2 は、サーバ装置 3 1 の同期制御部 1 2 からコマンド・データと処理開始通知情報を受信すると、受信したコマンド・データと処理開始通知情報を同期情報として蓄積する。

【0024】

また、同期制御部 1 2 は、処理時間管理部 1 4 からの処理終了通知情報を他のサーバ装置 3 2 の同期制御部 6 2 に送信する。サーバ装置 3 2 の同期制御部 6 2 は、処理終了通知情報を受信すると、同期情報として蓄積したコマンド・データおよび処理開始情報を同期制御部 6 2 から消去する。

40

【0025】

冗長制御部 1 3 は、ヘルスチェックによって他サーバ装置 3 2 の正常性および他サーバ装置 3 2 との伝送路の正常性を監視する。冗長制御部 1 3 は、伝送路を介して他のサーバ装置 3 2 の冗長制御部 6 3 に定期的にヘルスチェックデータを送信し、そのサーバ装置 3 2 の冗長制御部 6 3 からヘルスチェックデータを受信することによって、伝送路の状態を監視する。冗長制御部 1 3 は、監視結果からヘルスチェック送受信情報を生成して処理時

50

間管理部 1 4 に送信する。ヘルスチェック送受信情報は、ヘルスチェックデータを送信した時刻とヘルスチェックデータを受信した時刻を含む情報である。

【0026】

また、冗長制御部 1 3 は、自装置の動作状態およびインタフェースのリンク状態を監視する。冗長制御部 1 3 は、自装置で処理を継続できない障害を検出すると、他のサーバ装置 3 2 の冗長制御部 6 3 に障害情報を通知する。障害情報は、サーバ装置 3 1 の故障を他のサーバ装置 3 2 の冗長制御部 6 3 に通知するための情報である。同期制御部 6 2 は、冗長制御部 6 3 が障害情報を受信したことを検知すると、蓄積している同期情報を順序制御部 6 1 に送信する。順序制御部 6 1 は、同期制御部 6 2 から同期情報を受信すると、サーバ装置 3 1 の順序制御部 1 1 が行っていた処理を、同期情報を用いて引き継ぎ、その後の処理を開始する。 10

【0027】

処理時間管理部 1 4 は、伝送路の負荷状態とコマンド・データの処理時間を管理する。処理時間管理部 1 4 は、冗長制御部 1 3 からヘルスチェック送受信情報を受信すると、そのヘルスチェック送受信情報に含まれるヘルスチェックデータの送信時刻と受信時刻からヘルスチェックの所要時間を算出する。また、処理時間管理部 1 4 は、ヘルスチェックデータが送信された後、一定時間内に応答のヘルスチェックデータを受信されていなければ、そのサーバ装置 3 2 をコマンド・データの送信先から除外するよう順序制御部 1 1 に要求する。

【0028】

20

また、処理時間管理部 1 4 は、順序制御部 1 1 から処理開始通知情報と処理終了通知情報を受信し、それらを受信した時刻を記憶する。そして、処理時間管理部 1 4 は、処理開始通知情報を受信した時刻と、処理終了通知情報を受信した時刻とから実処理の所要時間を算出し、算出した実処理の所要時間によって、サーバ装置 3 1, 3 2 の負荷状態を監視する。

【0029】

処理時間監視部 1 4 は、コマンド・データの送信先から除外されていないサーバ装置 3 1, 3 2 の中から、ヘルスチェックの所要時間および実処理の所要時間を用いて、送信先とするサーバ装置を決定し、順序制御部 1 1 に通知する。また、処理開始通知情報を受信した後、それに対応する処理終了通知情報を所定時間内に受信できなければ、処理時間管理部 1 4 は順序制御部 1 1 にコマンド・データの再送を要求する。 30

【0030】

図 3 は、本実施形態のサーバ装置 3 1, 3 2 におけるヘルスチェックを行う際の動作を示すフローチャートである。冗長制御部 1 3 は、伝送路を介して接続されている他のサーバ装置 3 2 の冗長制御部 6 3 にヘルスチェックデータを送信する（ステップ 100）。処理時間管理部 1 4 は、そのサーバ装置 3 2 の冗長制御部 6 3 から所定の時間内にヘルスチェックデータを受信したか否かを判定する（ステップ 101）。 30

【0031】

サーバ装置 3 2 の冗長制御部 6 3 からヘルスチェックデータを所定の時間内に受信すると、処理時間管理部 1 4 は、ヘルスチェックの所要時間を算出する（ステップ 102）。 40
処理時間管理部 1 4 は算出したヘルスチェックの所要時間に基づいて送信先となるサーバ装置を指定する（ステップ 103）。サーバ装置 3 2 の冗長制御部 6 3 からヘルスチェックデータを所定の時間内に受信しないと、所要時間管理部 1 4 は、サーバ装置 3 2 を送信先から除外する（ステップ 104）。

【0032】

図 4 は、本実施形態のサーバ装置 3 1, 3 2 における実処理を行う際の動作を示すフローチャートである。順序制御部 1 1 は、コマンド・データを受信する（ステップ 150）。順序制御部 1 1 は、受信したコマンド・データを自装置のタスク処理部 5 1 または送信先となるサーバ装置 3 2 のタスク転送部 4 2 に転送し、その後、処理開始通知情報を処理時間管理部 1 4 に送信する（ステップ 151）。 50

【0033】

処理時間管理部14は、対応する処理終了通知情報を所定の時間内に受信したか否かを判定する（ステップ152）。対応する処理終了通知情報を所定の時間内に受信できなければ、処理時間管理部14は順序制御部11にコマンド・データの再送を要求する（ステップ153）。対応する処理終了通知情報を所定の時間内に受信すれば、処理時間管理部14は実処理の所要時間を算出する（ステップ154）。処理時間管理部14は、新たに算出した実処理の所要時間に基づいて送信先を指定する（ステップ155）。

【0034】

図5は、本実施形態の順序制御部11の構成を示すブロック図である。本順序制御部11は、バッファ格納処理部111、次コマンド・データ設定処理部112、コマンド・データ送信処理部113、処理結果受信処理部114、バッファ削除処理部115、および処理結果送信処理部116を有している。

【0035】

バッファ格納処理部111は、外部装置から受信したコマンド・データをメモリ上（不図示）に格納し、処理順序キューに登録する。

【0036】

次コマンド・データ設定処理部112は、処理結果受信処理部114から次の処理が実行可能となった旨の通知を受け取ると、バッファ格納処理部111に登録されたコマンド・データを読み込んで、コマンド・データ送信処理部113に送る。

【0037】

コマンド・データ送信処理部113は、処理時間管理部14から他のサーバ装置32が送信先として指定されていなければ、コマンド・データ設定処理部112からのコマンド・データをタスク処理部51に送る。また、コマンド・データ送信処理部113は、処理時間管理部14から他のサーバ装置32が送信先として指定されていれば、コマンド・データ設定処理部112からのコマンド・データを自装置のタスク処理部51または送信先となるサーバ装置32のタスク転送部42にコマンド・データを転送する。コマンド・データを転送した後、コマンド・データ送信処理部113は処理開始通知情報を処理時間管理部14に送る。

【0038】

処理結果受信処理部114は、タスク処理部51から処理結果を受け取ると、その処理結果を処理結果送信処理部116に送るとともに、処理終了通知情報を処理時間管理部14に送る。また、それとともに、処理結果受信処理部114は、次コマンド・データ設定処理部112に次の処理が実行可能となった旨を通知し、バッファ削除処理部115に処理が完了した旨を通知する。

【0039】

バッファ削除処理部115は、処理結果受信処理部114から処理の完了を通知されると、その処理が完了したコマンド・データをバッファから削除する。

【0040】

処理結果送信処理部116は、処理結果受信処理部114から処理結果を受けると、処理結果の順序を正常にして、外部装置に処理結果を返送する。

【0041】

図6は、本実施形態の順序制御部11の動作を示すフローチャートである。バッファ格納処理部111は、外部装置から受信したコマンド・データをメモリ上に格納するとともに、処理順序キューに登録する。

【0042】

次コマンド・データ設定処理部112は、バッファ格納処理部111の処理順序キューにコマンド・データが登録されているか否かを判定する（ステップ200）。バッファ格納処理部111に次のコマンド・データが登録されていると、次コマンド・データ設定処理部112は、バッファ格納処理部111のメモリ上からコマンド・データを読み込む（ステップ201）。

10

20

30

40

50

【0043】

コマンド・データ送信処理部113は、処理時間管理部14から他のサーバ装置32が送信先として指定されていなければ、読み込まれたコマンド・データを自装置のタスク処理部51に送る（ステップ202）。コマンド・データをタスク処理部51に送信した後、コマンド・データ送信処理部113は処理開始通知情報を処理時間管理部14に送る（ステップ203）。タスク処理部51は、処理が完了すると、処理結果を処理結果受信処理部114に送る（ステップ204）。

【0044】

処理結果受信処理部114は、タスク処理部51から処理結果を受け取ると、処理終了通知情報を処理時間管理部14に送る（ステップ205）。 10

【0045】

また、処理結果受信処理部114は、次コマンド・データ設定処理部112に次の処理が実行可能となった旨を通知する（ステップ206）。

【0046】

さらに、処理結果受信処理部114は、バッファ削除処理部115に処理が完了した旨を通知する。バッファ削除処理部115は、処理結果受信処理部114から処理が完了した旨を通知されると、その処理が完了したコマンド・データをバッファから削除する（ステップ207）。

【0047】

次に、処理結果受信処理部114は、処理結果送信処理部116に処理結果を送信する 20。処理結果送信処理部116は、処理結果受信処理部114から処理結果を受け取ると、順序を正常にして処理結果を外部装置に返送する（ステップ208）。

【0048】

図7は、本実施形態の同期制御部12の構成を示すブロック図である。同期制御部12は、コマンド・データ転送処理部121および処理状態転送処理部122を有している。

【0049】

コマンド・データ転送処理部121は、コマンド・データを他のサーバ装置32のタスク転送部42内にある同期制御部62に送信する。

【0050】

処理状態転送処理部122は、処理時間管理部14から処理開始通知情報と処理終了通知情報とを受信し、受信した処理開始通知情報と処理終了通知情報を同期制御部62に送信する。 30

【0051】

他のサーバ装置32では、タスク転送部42内の同期制御部62がコマンド・データ転送処理121と処理状態転送処理部122からコマンド・データおよび処理開始通知情報を受信して同期情報として蓄積する。また、タスク転送部42内の同期制御部62は、処理状態転送処理部122からの処理終了通知情報を受信すると、同期情報として蓄積したコマンド・データおよび処理開始通知情報を同期制御部62から消去する。

【0052】

図8は、本実施形態の冗長制御部13の構成を示すブロック図である。冗長制御部13 40は、ヘルスチェック監視処理部131および外部IF監視処理部132を有している。

【0053】

ヘルスチェック監視処理部131は、伝送路を介して他のサーバ装置32の冗長制御部63にヘルスチェックデータを送信し、そのサーバ装置32の冗長制御部63から応答のヘルスチェックデータを受信する。ヘルスチェック監視処理部131は、ヘルスチェックデータの送受信を示すヘルスチェック送受信情報を生成して処理時間管理部14に送信する。

【0054】

外部IF監視処理部132は、自装置（サーバ装置31）の動作状態およびインタフェースのリンク状態を監視する。外部IF監視処理部132は、物理レイヤレベルの状態を 50

監視してサーバ装置 3 1 の故障を検出する。サーバ装置 3 1 の故障を検出すると、外部 I/F 監視処理部 1 3 2 は障害情報を冗長制御部 6 3 に通知する。

【0055】

図 9 は、本実施形態の処理時間管理部 1 4 の構成を示すブロック図である。処理時間管理部 1 4 は、ヘルスチェック送受信情報受信部 1 4 1、ヘルスチェックタイムアウト監視部 1 4 2、送受信情報受信部 1 4 3、送信タイムアウト監視部 1 4 4、および処理負荷演算部 1 4 5 を有している。

【0056】

ヘルスチェック送受信情報受信部 1 4 1 は、冗長制御部 1 3 からヘルスチェック送受信情報を受信すると、ヘルスチェックタイムアウト監視部 1 4 2 と処理負荷演算部 1 4 5 にヘルスチェック送受信情報を送信する。 10

【0057】

ヘルスチェックタイムアウト監視部 1 4 2 は、ヘルスチェック送受信情報受信部 1 4 1 からヘルスチェックデータの送信を示すヘルスチェック送受信情報を受信した後、所定時間内にヘルスチェックデータの受信を示すヘルスチェック送受信情報を受信できたか否かを判定する。ヘルスチェックタイムアウト監視部 1 4 2 は、ヘルスチェックが所定時間内に完了しなければ、そのサーバ装置 3 2 を送信先から除外するよう順序制御部 1 1 に要求する。

【0058】

送受信情報受信部 1 4 3 は、順序制御部 1 1 から処理開始通知情報または処理終了通知情報を受信すると、送信タイムアウト監視部 1 4 4、処理負荷演算部 1 4 5、および同期制御部 1 2 に処理開始通知情報または処理終了通知情報を送信する。 20

【0059】

送信タイムアウト監視部 1 4 4 は、処理開始通知情報を受信した時刻から所定時間内に処理終了通知情報を受信したか否かを判定する。送信タイムアウト監視部 1 4 4 は、所定時間内に処理終了通知情報を受信できなければ、順序制御部 1 1 にコマンド・データを再送するように要求する。

【0060】

処理負荷演算部 1 4 5 は、ヘルスチェック送受信情報受信部 1 4 1 からヘルスチェック送受信情報を受信すると、そのヘルスチェック送受信情報に含まれる送信時刻と受信時刻からヘルスチェックの所要時間を算出する。 30

【0061】

また、処理負荷演算部 1 4 5 は、送受信情報受信部 1 4 3 から処理開始通知情報と処理終了通知情報を受信すると、それらを受信した時刻を記憶する。そして、処理負荷演算部 1 4 5 は、処理開始通知情報を受信した時刻と、処理終了通知情報を受信した時刻とから処理開始から処理完了までの実処理の所要時間を算出する。なお、送受信情報受信部 1 4 3 から処理負荷分散演算部 1 4 5 へは自装置の処理開始通知情報および処理終了通知情報と、他のサーバ装置 3 2 の処理開始通知情報および処理終了通知情報とが通知される。そのため、処理負荷演算部 1 4 5 は、自装置の実処理の所要時間と他のサーバ装置 3 2 の実処理の所要時間とを算出することとなる。 40

【0062】

処理負荷演算部 1 4 5 は、算出したヘルスチェックの所要時間と実処理の所要時間とから、送信先とするサーバ装置を決定し、順序制御部 1 1 に通知する。例えば、処理負荷演算部 1 4 5 は、ヘルスチェックの所要時間が最も短いサーバ装置を送信先として順序制御部 1 1 に通知する。順序制御部 1 1 は、送信先とされたサーバ装置にコマンド・データを転送する。また、処理負荷演算部 1 4 5 は、実処理の所要時間を算出し、算出した実処理の所要時間が最も短いサーバ装置を送信先として順序制御部 1 1 に通知する。

【0063】

以上説明したように、本実施形態によれば、サーバ装置 3 1 のタスク転送部 4 1 がタスク処理部 5 1 の実処理の負荷および他サーバ装置 3 2 の実処理の負荷に加えて、他サーバ 50

装置 3 2 との間の伝送路の負荷を監視し、それら実処理の負荷と伝送路の負荷に基づいてコマンド・データの割り当て先を決定する。そのため、実処理だけでなく伝送路の負荷をも考慮してネットワークに接続された負荷を分散することができ、応答の即時性を向上させることができる。

【0064】

なお、本実施形態ではコマンド・データを一旦転送した後に、そのコマンド・データの転送先を変更することはないが、本発明はこれに限定されるものではない。他の例として、サーバ装置 3 1 は、サーバ装置 3 2 にコマンド・データを転送してから一定時間内にサーバ装置 3 2 から受け付けを通知する応答が無ければ、このコマンド・データをサーバ装置 3 2 以外のサーバ装置（不図示）に転送して処理を行わせることにしてもよい。

10

【図面の簡単な説明】

【0065】

【図 1】本実施形態における処理の負荷分散装置の構成を示すブロック図である。

【図 2】本実施形態のサーバ装置 3 1, 3 2 の冗長構成を示すブロック図である。

【図 3】本実施形態のサーバ装置 3 1, 3 2 におけるヘルスチェックを行う際の動作を示すフローチャートである。

【図 4】本実施形態のサーバ装置 3 1, 3 2 における実処理を行う際の動作を示すフローチャートである。

【図 5】本実施形態の順序制御部 1 1 の構成を示すブロック図である。

【図 6】本実施形態の順序制御部 1 1 の動作を示すフローチャートである。

20

【図 7】本実施形態の同期制御部 1 2 の構成を示すブロック図である。

【図 8】本実施形態の冗長制御部 1 3 の構成を示すブロック図である。

【図 9】本実施形態の処理時間管理部 1 4 の構成を示すブロック図である。

【符号の説明】

【0066】

1 ノード装置

2 伝送路装置

3 サーバ装置

4 タスク転送部

5 タスク処理部

30

1 1 順序制御部

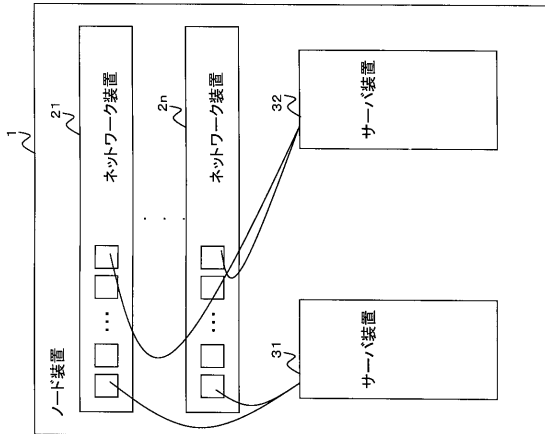
1 2 同期制御部

1 3 冗長制御部

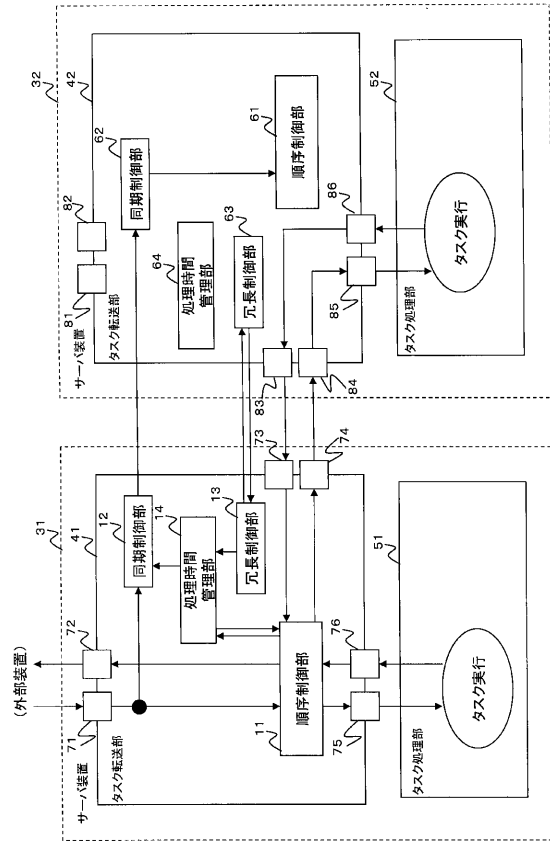
1 4 処理時間管理部

1 0 0 ～ 1 0 4、1 5 0 ～ 1 5 5、2 0 0 ～ 2 0 9 ステップ

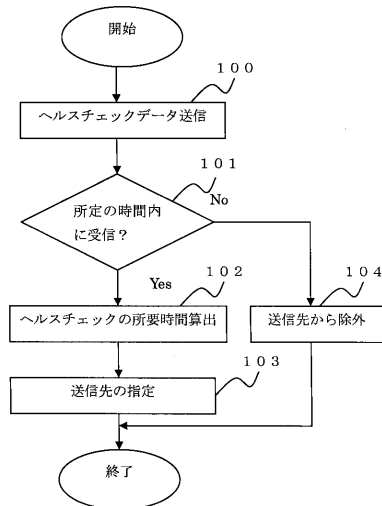
【図 1】



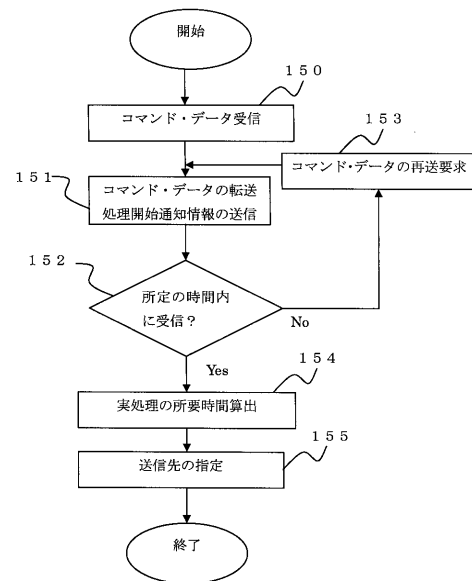
【図 2】



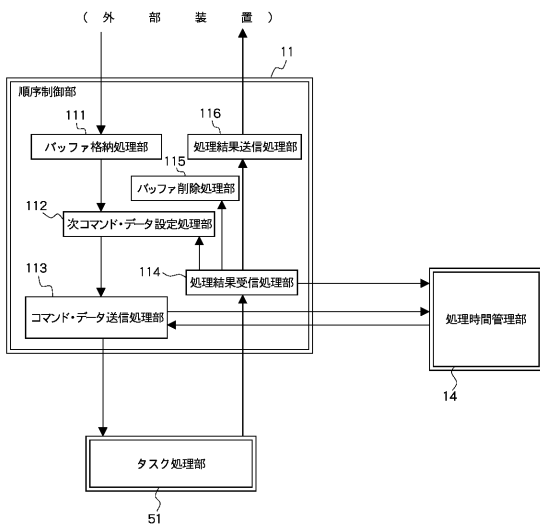
【図 3】



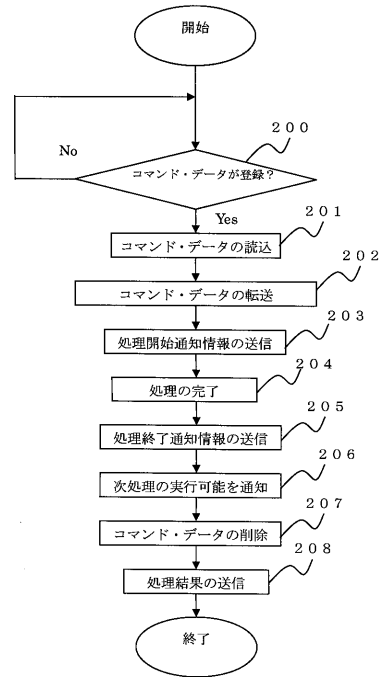
【図 4】



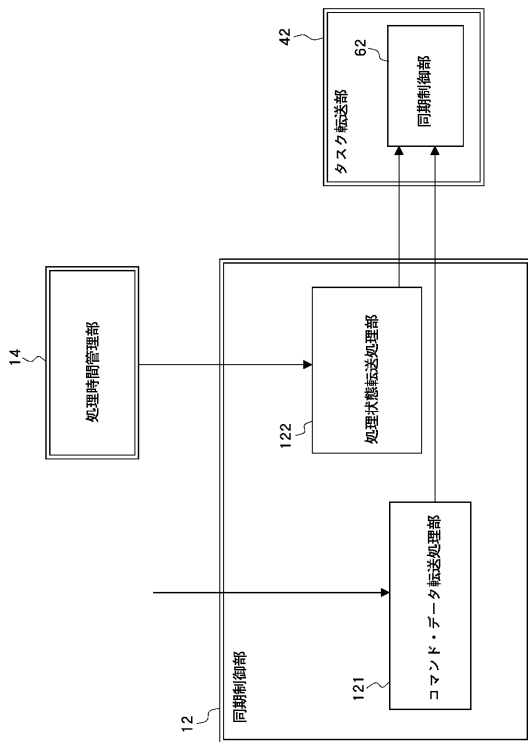
【図 5】



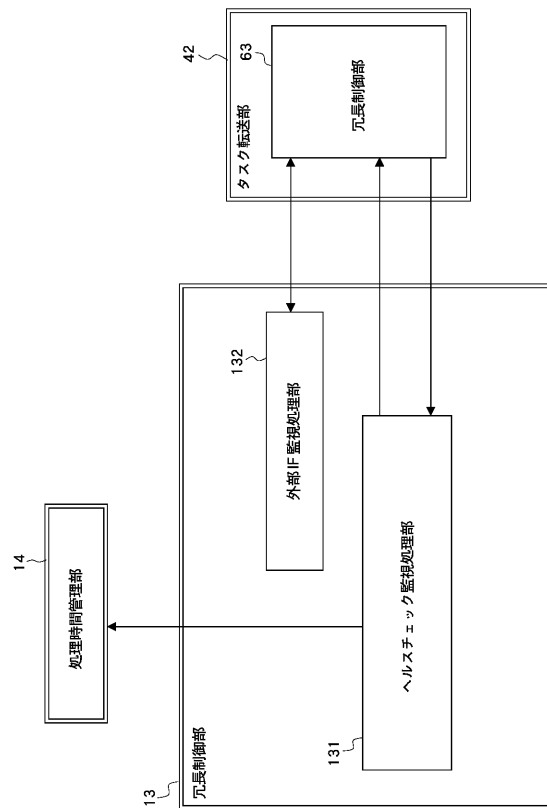
【図 6】



【図 7】



【図 8】



【図 9】

