(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2012 年 5 月 3 日(03.05.2012)

PCT

(10) 国際公開番号
WO 2012/056487 A1

(51) 国際特許分類:
G06F 11/18 (2006.01) *G06F 9/52* (2006.01)

(21) 国際出願番号: PCT/JP2010/006290

(22) 国際出願日: 2010 年 10 月 25 日(25.10.2010)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人 (米国を除く全ての指定国について): 株式会社日立製作所 **(HITACHI, LTD.)** [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目６番６号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 金　成昊 **(KIM, Sungho)** [KR/JP]; 〒2440817 神奈川県横浜市戸塚区吉田町２９２番地　株式会社日立製作所　システム開発研究所内 Kanagawa (JP). 西島英児 **(NISHIJIMA, Eiji)** [JP/JP]; 〒2440817 神奈川県横浜市戸塚区吉田町２９２番地　株式会社日立製作所　システム開発研究所内 Kanagawa (JP).

(74) 代理人: 井上　学, 外 **(INOUE, Manabu** et al.); 〒1008220 東京都千代田区丸の内一丁目６番１号 株式会社日立製作所内 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

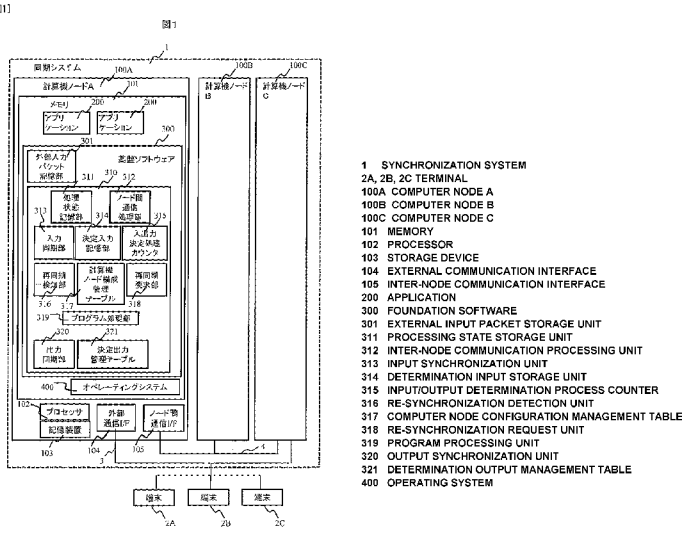
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[続葉有]

(54) **Title**: COMPUTER SYSTEM

(54) 発明の名称 : 計算機システム

[図1]



(57) **Abstract**: Provided is a computer system having multiple computer nodes, with each computer node having: an external communication unit that receives processing requests; an application unit that executes processes in response to said processing requests; a synchronization unit that synchronizes the processes executed by each computer node with those of the other computer nodes; and an inter-node communication unit that transmits and receives information with respect to the other computer nodes. The synchronization unit transmits processing requests received by the external communication unit to the other computer nodes via the inter-node communication unit. Furthermore, the synchronization unit receives processing requests from the other computer nodes via the inter-node communication unit. Based on the number of computer nodes that have received the same processing request through the external communication unit, the synchronization unit selects a processing request to be executed by the application unit, and the application unit executes the selected processing request.

(57) 要約:

[続葉有]

WO 2012/056487 A1

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

複数の計算機ノードを有する計算機システムであって、各計算機ノードは、処理要求を受信する外部通信部と、当該処理要求に従って処理を実行するアプリケーション部と、各計算機ノードで実行される処理を他の計算機ノードとの間で同期させる同期部と、他の計算機ノードとの間で情報の送受信を実行するノード間通信部とを有する。同期部は、外部通信部が受信した処理要求をノード間通信部を介して他の計算機ノードに送信する。また、同期部は、ノード間通信部を介して他の計算機ノードからも処理要求を受信する。そして同期部は、同じ処理要求を外部通信部を経由して受信した計算機ノードの数に基づいて、アプリケーション部に実行させる処理要求を選択し、アプリケーション部は選択された処理要求を実行する。

明　細　書

発明の名称 ： 計算機システム

技術分野

[0001] 複数の計算機がネットワークで接続された多重システムにおいて計算機間で同期をとる技術に関する。

背景技術

[0002] ネットワークで接続されている複数の計算機が各々同じ業務アプリケーションを実行することで、高信頼及び高可用なシステムを構築する技術がある。特に、障害時の業務の連続性を保証するために、多数の計算機を同期させて運用するフォールトトレランスコンピュータ（以下、ＦＴＣと呼ぶ。）技術が幅広く使用されている。FTCには、CPUのバスクロックレベルでLOCKSTEPを実現する同期方式が幅広く使用されているが（特許文献１参照）、近年CPUの構成の複雑化につれ、ソフトウェアレベルでの同期方式が注目されている（特許文献２参照）。

先行技術文献

特許文献

[0003] 特許文献1：特開2006-209565

特許文献2：特開2009-217503

発明の概要

発明が解決しようとする課題

[0004] 従来のソフトウェアレベルでの同期方式では、同期ポイントをアプリケーション内部に設けて同期を取っているため、同期処理にオーバヘッドが発生するという課題がある。また、フォルトトラレンスコンピュータの出力処理においては、ＦＴＣを構成する複数の計算機各々からの出力を用いて多数決を行うことでＦＴＣからの出力値を決定するボーターという装置が別途必要である。高信頼及び高可用性を保つためには、ボーター装置も多重化しなければならない。

[0005] そこで、ボータ装置の様な装置を別途用いることなく複数計算機間でソフトウェアレベルでの同期を実現する技術を提供する。また、低オーバヘッドな同期方法も提供する。

課題を解決するための手段

[0006] 複数の計算機ノードを有する計算機システムであって、各計算機ノードは、処理要求を受信する外部通信部と、当該処理要求に従って処理を実行するアプリケーション部と、各計算機ノードのアプリケーション部で実行される処理の他の計算機ノードとの間での同期を制御する同期部と、他の計算機ノードとの間で情報の送受信を実行するノード間通信部とを有する。各計算機ノードの同期部は、外部通信部が受信した処理要求識別情報が付与された処理要求に、当該計算機ノードの識別情報を送信元計算機ノード識別情報として付与して前記ノード間通信部を介して他の計算機ノードに送信する。また、各計算機ノードの同期部は、ノード間通信部を介して他の計算機ノードから当該他の計算機ノードの識別情報が付与された処理要求を受信する。そして、各計算機ノードの同期部は、同じ処理要求識別情報が付与された処理要求を外部通信部を経由して受信した計算機ノードの数に基づく第一の所定の判断基準に従って、ノード間通信部が他の計算機ノードから受信した処理要求と前記外部通信部が受信した処理要求の中から、アプリケーション部に実行させる処理要求を選択し、アプリケーション部は選択された処理要求を実行する。

発明の効果

[0007] 低オーバヘッドな同期処理が可能なフォルトトラレンスコンピュータを実現することができる。また、多数決判定のための特別な装置であるボーターを必要とせず、システム構成がシンプルなフォルトトレランスコンピュータの構築が可能となる。

図面の簡単な説明

[0008] [図1]多数決同期システムの全体構成の一例を示す図である。

[図2]外部入力パケットの構成の一例を示す図である。

[図3]外部入力パケット管理テーブルの構成の一例を示す図である。

[図4]計算機ノード処理状態テーブルの構成の一例を示す図である。

[図5]計算機ノードの処理状態遷移の一例を示す図である。

[図6]ノード間通信パケットの構成の一例を示す図である。

[図7]計算機ノード構成管理テーブルの構成の一例を示す図である。

[図8]決定入力管理テーブルの構成の一例を示す図である。

[図9]決定済み入力管理テーブルの構成の一例を示す図である。

[図10]決定出力管理テーブルの構成の一例を示す図である。

[図11]基盤ソフトウェア処理の一例を示すフローチャートである。

[図12]処理対象入力パケット選定開始確認処理の一例を示すフローチャートである。

[図13]処理対象入力パケット選定処理の一例を示すフローチャートである。

[図14]入力決定パケット送信処理の一例を示すフローチャートである。

[図15]入力処理状態一致処理の一例を示すフローチャートである。

[図16]入力多数決処理の一例を示すフローチャートである。

[図17]選定入力パケット処理の一例を示すフローチャートである。

[図18]実行結果決定処理の一例を示すフローチャートである。

[図19]出力処理状態一致処理の一例を示すフローチャートである。

[図20]出力多数決処理の一例を示すフローチャートである。

[図21]リーダー選出処理の一例を示すフローチャートである。

[図22]リーダー処理状態一致化処理の一例を示すフローチャートである。

[図23]ボーター処理の一例を示すフローチャートである。

[図24]出力異常処理の一例を示すフローチャートである。

[図25]再同期検知処理および再同期処理の一例を示すフローチャートである。

[図26]再同期検知処理および再同期処理において用いられるノード間通信パケットの一例を示す図である。

[図27]ケースＡにおける、各計算機ノードの処理シーケンスの一例を示す図

である。

[図28]ケースＡににおける、各計算機ノードの外部入力パケット管理テーブル、計算機ノード処理状態テーブル、入出力決定処理カウンタの内容を時間別に例示した図である。

[図29]ケースＡにおける、送信に対する各計算機ノードのノード間通信パケットの内容を時間別に例示した図である。

[図30]ケースＡにおける、リスト作成に対する各計算機ノードのノード間通信パケットの内容を時間別に例示した図である。

[図31]ケースＡにおける、任意の計算機ノードの計算機ノード構成管理テーブルの内容を時間別に例示した図である。

[図32]ケースＡにおける、各計算機ノードの決定入力管理テーブルの内容の一例を示す図である。

[図33]ケースＡにおける、計算機ノード１００Ｂの決定済み入力管理テーブルの内容の一例を示す図である。

[図34]ケースＢにおける、各計算機ノードの処理シーケンスの一例を示す図である。

[図35]ケースＢにおける、各計算機ノードの外部入力パケット管理テーブル、計算機ノード処理状態テーブル、入出力決定処理カウンタの内容を時間別に例示した図である。

[図36]ケースＢにおける、送信に対する各計算機ノードのノード間通信パケットの内容を時間別に例示した図である。

[図37]ケースＢにおける、リスト作成に対する各計算機ノードのノード間通信パケットの内容を時間別に示す図である。

[図38]ケースＢにおける、計算機ノード１００Ａ、１００Ｂの計算機ノード構成管理テーブルの内容を時間別に例示した図である。

[図39]ケースＢにおける、計算機ノード１００Ｃの計算機ノード構成管理テーブルの内容を時間別に示す図である。

[図40]ケースＢにおける、各計算機ノードの決定入力管理テーブルの内容の

一例を示す図である。

[図41]ケースＢにおける、計算機ノード１００Ｂの決定済み入力管理テーブルの内容の一例を示す図である。

[図42]ケースＣにおける、各計算機ノードの処理シーケンスの一例を示す図である。

[図43]ケースＣにおける、時間Ｔ１間およびＴ１'間の送信に対する各計算機ノードのノード間通信パケットの内容の一例を示す図である。

[図44]ケースＣにおける、時間Ｔ２間の送信についての各計算機ノードのノード間通信パケットの内容の一例を示す図である。

[図45]ケースＣにおける、時間Ｔ３間の送信についての各計算機ノードのノード間通信パケットの内容の一例を示す図である。

[図46]ケースＣにおける、時間Ｔ１間のリスト作成についての計算機ノード１００Ａ、１００Ｂのノード間通信パケットの内容の一例を示す図である。

[図47]ケースＣにおける、時間Ｔ２間のリスト作成についての各計算機ノードのノード間通信パケットの内容の一例を示す図である。

[図48]ケースＣにおける、任意の計算機ノードの計算機ノード構成管理テーブルの内容を時間別に示す図である。

[図49]ケースＣにおける、各計算機ノードの決定出力管理テーブルの内容の一例を示す図である。

[図50]新規計算機ノードを追加する場合のシステムの全体構成例を示す図である。

発明を実施するための形態

[0009] 以下、実施の形態を説明する。説明は実施例１と実施例２とに分けて行う。実施例１では実施形態の詳細について説明を行った後、更に３つの具体的なケースを想定してケース毎に実施態様を説明する。実施例２では、実施例１で説明した内容を踏まえて、多数決同期システムに新たな計算機ノードを加える場合における処理を説明する。尚、図面においては、同じ要素には同じ参照番号を用いる。

[0010] 本発明が実施の形態に制限されることは無く、本発明の思想に合致するあらゆる応用例が本発明に含まれる。

実施例 1

[0011] ＜全体構成（図１）＞

図１は多数決同期システムの全体構成の一例を示す図である。多数決同期システムは一つの同期システム１（フォルトトレランスコンピュータとも呼ぶ）と複数の端末２および外部ネットワーク３により構成される。同期システム１は複数台の計算機ノード１００と内部ネットワーク４により構成されている。

同期システム１は外部ネットワーク３を経由して端末２から外部入力パケット５００を受信する。端末２から送信される外部入力パケット５００の内、因果関係のある複数の外部入力パケット５００はその送信順序が守られて同期システム１に到着する。なぜならば、このような外部入力パケット５００については、同期システム１が外部入力パケット５００を受信した後、当該外部入力パケット５００に対する応答を同期システム１から端末２が受け取った後にはじめて、次の外部入力パケット５００が端末２から同期システム１に送信されるからである。その他の外部入力パケット５００については、端末２における送信順序と同期システム１における受信順序は必ずしも一致しない。従って、順序を守って処理する必要がある外部入力パケット５００に関しては、外部入力パケット５００に順序を示す情報（タイムスタンプやシーケンス番号）が付与されている。

尚、外部入力パケット５００は端末２から同期システム１へマルチキャスト方式で送信され、これに対する応答（処理結果等）は同期システム１から端末２へマルチキャスト方式で返信される。また、同期システム１で実行されるアプリケーションは内部状態（オープン状態のファイル、ネットワークコネクション等）を持たないものとする。

[0012] 計算機ノード１００はメモリ１０１、プロセッサ１０２、二次記憶装置１０３、外部通信インタフェース１０４（図中、インタフェースをI／Fと略記

）、ノード間通信Ｉ／Ｆ１０５のハードウェアを備え、これらが内部バス等の通信路によって接続されている。計算ノード１００を制御するソフトウェアプログラムは、アプリケーション２００、基盤ソフトウェア３００、オペレーティングシステム４００から構成されている。さらに基盤ソフトウェア３００は外部入力パケット記憶部３０１と入出力同期部３１０に大別され、入出力同期部３１０は処理状態記憶部３１１、ノード間通信処理部３１２、入力同期部３１３、決定入力記憶部３１４、入出力決定処理カウンタ３１５、再同期検知部３１６、計算機ノード構成管理テーブル３１７、再同期要求部３１８、プログラム処理部３１９、出力同期部３２０、決定出力管理テーブル３２１から構成される。同期システム１内の複数の計算機ノード１００は、いずれも同様のハードウェア／ソフトウェア構成を有する。図１では３台の端末2A、２B、２Ｃと３台の計算機ノード１００A、１００B、１００Cの例が示されているが、特に台数は限定されない。またアプリケーション２００の個数や種類の数は限定されるものでは無い。

[0013] 端末２と計算機ノード１００は外部ネットワーク３を介して接続される。計算機ノード１００側では、外部ネットワーク３はノード間通信Ｉ／Ｆ１０５に接続される。同期システム１内の各計算機ノード１００A、１００B、１００Cは、ノード間通信Ｉ／Ｆに接続された内部ネットワーク４を介して接続される。

[0014] 端末２は、外部ネットワーク３を通じて同期システム１に要求を送信し、同期システム１が当該要求を処理して結果を端末２に返信する。同期システム１内の各計算機ノード１００は、外部ネットワーク３及び外部通信I／Ｆ１０４を通じて受信した端末２からの要求を、内部ネットワーク４とノード間通信Ｉ／Ｆ１０５を介して他の計算機ノード１００にも送信し、他の計算機ノードと協調して動作して処理し、処理結果を端末２に返信する。

[0015] 計算機ノード１００内の二次記憶装置１０３はアプリケーション２００、基盤ソフトウェア３００、オペレーティングシステム４００を格納している。プロセッサ１０２は二次記憶装置１０３からアプリケーション２００、基

盤ソフトウェア３００、オペレーティングシステム４００をメモリ１０１への読み込んで展開し、基盤ソフトウェア３００やアプリケーション２００、オペレーションシステム４００を実行すると共に、外部通信I／Ｆ１０４やノード間通信Ｉ／Ｆ１０５からの割込みも処理する。以降、アプリケーション２００やオペレーションシステム４００や基盤ソフトウェア３００（その構成要素を含む）が実行する処理として説明される動作は、実際にはプロセッサ１０２がこれらのソフトウェアを実行することによって行われる。外部通信I／Ｆ１０４は外部ネットワーク３を通じて端末２とのパケットの送受信を行う。ノード間通信Ｉ／Ｆ１０５は内部ネットワーク４を通じて計算機ノード１００間のパケットの送受信を行う。

[0016] メモリ１０１内に展開されたオペレーティングシステム４００は基盤ソフトウェア３００やアプリケーション２００の動作を制御する。基盤ソフトウェア３００の外部入力パケット記憶部３０１と入出力同期部３１０は異なるタスクとして動作し、アプリケーション２００は必要に応じてプログラム処理部３１９から実行される。

[0017] 外部入力パケット記憶部３０１は外部入力パケット管理テーブル５５０を保有し、外部通信I／Ｆ１０４が受信した外部入力パケット５００を外部入力パケット管理テーブル５５０に格納する。外部入力パケット５００の詳細は図２と共に後述する。外部入力パケット管理テーブル５５０の詳細は図３と共に後述する。

処理状態記憶部３１１は計算機ノード処理状態テーブル７００を保有し、計算機ノード１００の処理状態が記録される。計算機ノード処理状態テーブル７００の詳細は図４と共に後述する。

[0018] ノード間通信処理部３１２は計算機ノード１００間の通信処理を制御し、ノード間通信パケット６００の送受信する。ノード間通信パケット６００の詳細は図１０と共に後述する。

[0019] 入力同期部３１３は処理対象入力パケット選定開始処理Ｓ１１３および処理対象入力パケット選定処理Ｓ１１４を実効する。処理対象入力パケット選

定開始処理Ｓ１１３の詳細は図１２と共に後述する。処理対象入力パケット選定処理Ｓ１１４の詳細は図１３～図１６と共に後述する。

[0020] 決定入力記憶部３１４は決定入力管理テーブル８００および決定済み入力管理テーブル８１０を保有し、入力同期部３１３にて決定され、入出力同期部３１０が処理する外部入力パケット５００の情報を保持する。決定入力管理テーブル８００の詳細は図７と共に後述する。決定済み入力管理テーブル８１０の詳細は図８と共に後述する。

[0021] 入出力決定処理カウンタ３１５は計算機ノード１００が実行する処理段階のカウンタ値を保持する。他の計算機ノード１００との間で入出力決定処理カウンタ３１５を比較することで、計算機ノード１００が実行する処理の周期ずれを識別する。

[0022] 再同期検知部３１６は複数の計算機ノード１００が協調動作している同期システム１に新しく参加しようとする計算機ノード１００や、一度協調動作している同期システム１から排除されたが再参加を試みている計算機ノード１００を検出して、受入れ処理を行う。再同期検知部３１６は再同期検知処理Ｓ１１１を実行する。再同期検知処理Ｓ１１１の詳細は図２５と共に後述する。

[0023] 計算機ノード構成管理テーブル３１７は処理対象入力パケット選定処理Ｓ１１４を実行し、選定入力パケット処理Ｓ１１５の処理対象となる外部入力パケット５００の決定に寄与する計算機ノード１００情報を保持するテーブルである。計算機ノード構成管理テーブル３１７の詳細は図６と共に後述する。

[0024] 再同期要求部３１８は複数の計算機ノード１００が協調動作している同期システム１に新しく参加しようとする計算機ノード１００や、一度協調動作している同期システム１から排除されたが再参加を試みている計算機ノード１００が実行する処理である。サイ同期要求部３１８は再同期処理Ｓ１１８を実行する。再同期処理Ｓ１１８の詳細は図２５と共に後述する。

[0025] プログラム処理部３１９は端末２からの処理要求に応じてアプリケーショ

ン２００を選択、起動を実行する処理である。選定入力パケット処理Ｓ１１５を実行する。選定入力パケット処理Ｓ１１５の詳細は図１７と共に後述する。

[0026] 出力同期部３２０は、各計算機ノード１００のアプリケーション２００の実行結果を用いて多数決判定を行い、各計算機ノード１００についてリーダーとモニターの役割分けを行い、端末２への実行結果送信処理を行う。出力同期部３２０は実行結果決定処理Ｓ１１６を実行する。実行結果決定処理Ｓ１１６の詳細は図１８～図２４と共に後述する。

[0027] 決定出力管理テーブル３２１は実行結果決定処理Ｓ１１６に寄与した計算機ノード１００の情報を保持するテーブルである。決定出力管理テーブル３２１の詳細は図９と共に後述する。

＜外部入力パケット（図２）＞

図２は外部入力パケット５００の構成の一例を表す図である。端末２は同期システム１へ外部入力パケット５００を送信することで要求命令を伝える。

[0028] 送信元情報５０１は送信元端末２を識別する情報である。シーケンス番号５０２は送信元端末２から発せられた複数の外部入力パケット５００各々を一意に識別する情報である。シーケンス番号５０２の値は、送信元端末２の識別子を含んだ値にすることによって、端末２間で重複しないようにすることができる。送信タイムスタンプ５０３は送信元端末２が外部入力パケット５００を発信した時刻である。順列処理要否フラグ５０４は送信元端末２が送信した複数の外部入力パケット５００間に一連の順序性が存在することを識別する情報である。順列処理ブロック数５０５は前述の一連の順序性が存在する複数の外部入力パケット５００の個数を示す。順列処理開始番号５０６は前述の一連の順序性が存在する複数の外部入力パケット５００の内、先頭に位置する外部入力パケット５００のシーケンス番号５０２である。コンテンツ５０７はアプリケーション２００の実行に関係する情報であるアプリ識別子５０８とパラメタ５０９から構成される。アプリ識別子５０８は計算

機ノード１００内で実行されるアプリケーション２００を指定する情報であり、パラメタ５０９にはアプリケーション２００に引渡されるパラメタ情報である。

＜外部入力パケット管理テーブル（図３）＞

図３は外部入力パケット管理テーブル５５０の構成の一例を示す図である。外部入力パケット管理テーブル５５０は外部入力パケット記憶部３０１内に保持され、計算機ノード１００が受信した外部入力パケット５００を格納するテーブルである。外部入力パケット管理テーブル５５０に格納される外部入力パケット５００数は特に限定されるものではない。

[0029] 番号５５１は格納した外部入力パケット５００を一意に特定する管理番号である。外部入力パケット管理テーブル５５０には、受信した外部入力パケット５００が改変されることなく格納されるため、番号５５１以外の列（要素）は外部入力パケット５００の要素と同様である。

＜計算機ノード処理状態テーブル（図４）＞

図４は計算機ノード処理状態テーブル７００の構成例を示す図である。計算機ノード処理状態テーブル７００は処理状態記憶部３１１内に保持され、計算機ノード１００の実行状態情報を格納する。計算機ノード処理状態テーブル７００は処理状態識別子７０１、付加情報７０２、Penalty７０３、処理状態７０４の列から構成される。

[0030] 処理状態識別子７０１と処理状態７０４は常に有効な値を保持する列であるが、付加情報７０２とPenalty７０３は処理状態識別子７０１に格納される値により有効な場合と無効な場合がある。

[0031] 処理状態識別子７０１は計算機ノード１００の処理状態を識別する値である。処理状態識別子として「REJOIN-WAIT」、「REJOIN-REQ」、「REJOIN-ACK」、「JOINED」、「INPUT」、「OUTPUT」がある。

[0032] 「REJOIN-WAIT」は当該計算機ノード１００が他の計算機ノード１００からの再同期要求を待っている状態を表す。この状態の時、処理状態７０４にも「REJOIN-WAIT」が設定されるが、付加情報７０２とPenalty７０３は有効な

値を保有しない（行７０５）。

[0033] 「REJOIN-REQ」は当該計算機ノード１００が再同期要求を行っている状態を表す。この状態の時、処理状態７０４にも「REJOIN-REQ」が設定されるが、付加情報７０２とPenalty７０３は有効な値を保有しない。（行７０６）

「REJOIN-ACK」は当該計算機ノード１００が他の計算機ノード１００からの再同期要求を受領した状態を表す。この状態の時、処理状態７０４にも「REJOIN-ACK」が設定されるが、付加情報７０２とPenalty７０３は有効な値を保有しない。（行７０７）

「JOINED」は当該計算機ノード１００が行った再同期要求に対して他の計算機ノード１００からの返答を取得し、取得した情報を当該計算機ノード１００に反映した状態を表す。この状態の時、処理状態７０４にも「JOINED」が設定されるが、付加情報７０２とPenalty７０３は有効な値を保有しない。（行７０８）

「INPUT」は当該計算機ノード１００が「入力同期」を行っている状態を表す。「入力同期」とは、計算機ノード１００の外部入力パケット管理テーブル５５０に格納されている外部入力パケット５００をアプリケーション２００で処理するためにおこなわれる処理である。具体的には、各計算機ノード１００間で多数決判定を行い、処理対象の外部入力パケット５００と、アプリケーション２００を実行する計算機ノード１００を決定する処理である。この状態の時、付加情報７０２とPenalty７０３が有効になる。付加情報７０２には入力同期ステップ７１０情報が格納される。入力同期ステップ７１０は、外部入力パケット管理テーブル５５０から取得する外部入力パケット５００の個数である。Penalty７０３は、例えば初期値として「０」、初期値以外の値として「１」のいずれかをとることができ、「１」である場合には当該計算機ノード１００が他の計算機ノード１００との間で行われる入力同期処理に一時不参加であったことを表す。処理状態７０４には「BEGIN」もしくは「END」が保持される。「BEGIN」は当該計算機ノード１００が「入力同期」を実行している状態を表し、「END」は「入力同期」が完了した状態を表す

（行７０９）。

[0034] 「OUTPUT」は当該計算機ノード１００が「出力同期」を行っている状態を表す。「出力同期」とは、他の計算機ノード１００との間でアプリケーション２００の処理結果を用いて多数決判定を行い、処理結果を確定すると共に、端末２へ結果を出力する役割を担う「リーダー」計算機ノード１００と「リーダー」の結果出力状態を監視する役の「モニター」計算機ノード１００とを決定して、端末２へアプリケーション２００の処理結果を出力する処理である。この状態の時、付加情報７０２は有効な値となり、「OUTPUT」、「LEADER」、「VOTER」の何れかの値を保持する。

[0035] 付加情報７０２が「OUTPUT」の時は、当該計算機ノード１００はアプリケーション２００の処理結果を他の計算機ノード１００との間で共有する処理を行う段階であることを表し、共有処理が実行中の間は処理状態７０４に「BEGIN」が保持され、共有処理が完了すると処理状態７０４に「END」が保持される。またPenalty７０３は、例えば初期値として「０」、初期値以外の値として「１」をとる有効な列となり、「１」である場合には当該計算機ノード１００が出力同期処理に一時不参加であったことを表す。

[0036] 付加情報７０２が「LEADER」の時は、当該計算機ノード１００は「リーダー」役と「モニター」役の計算機ノード１００の決定処理及び他の計算機ノード１００との間で決定結果を共有する処理を実行する状態であることを表し、処理を実行中の間は処理状態７０４に「BEGIN」が保持され、役割決定処理が完了すると処理状態７０４に「END」が保持される。

[0037] 付加情報７０２が「VOTER」の時は、当該計算機ノード１００は「リーダー」または「モニター」役としてアプリケーション２００の処理結果を出力する処理を行う状態であることを表し、「リーダー」役で出力処理を行う場合は処理状態７０４に「LEADER」が保持され、「モニター」役で出力処理を行う場合は処理状態７０４に「MONITOR」が保持される（行７１１）。

＜計算機ノード処理状態遷移図（図５）＞

図５は計算機ノード１００の処理状態遷移の一例を示す図である。図５中

のREJOIN-WAIT７０５、REJOIN-REQ７０６、REJOIN-ACK７０７、JOINED７０８、INPUT７０９、OUTPUT７１１は計算機ノード処理状態テーブル７００の処理状態識別子７０１に設定される識別子と一致する。

[0038] REJOIN-WAIT７０６処理状態の計算機ノード１００は、REJOIN-ACK７０７処理状態に遷移した後INPUT７０９処理状態に遷移する、もしくは直接INPUT７０９処理状態に遷移する。REJOIN-REQ７０５処理状態の計算機ノード１００はJOINED７０８処理状態に遷移した後、INPUT７０９処理状態に遷移する。INPUT７０９処理状態の計算機ノード１００はOUTPUT７１１処理状態もしくはREJOIN-REQ７０５処理状態に遷移する。OUTPUT７１１処理状態の計算機ノード１００はREJOIN-WAIT７０６処理状態もしくはREJOIN-REQ７０５処理状態に遷移する。

＜ノード間通信パケット（図６）＞

図６はノード間通信パケット６００の構成の一例を示す図である。ノード間通信パケット６００は計算機ノード１００の間で情報通信を行う際に使用される通信パケットである。ノード間通信パケット６００は計算機ノード識別子６０１、入出力決定処理カウンタ６０２、計算機ノード処理状態６０３、計算機ノード処理状態個別詳細情報６０４から構成される。

[0039] 計算機ノード識別子６０１は同期システム１内で計算機ノード１００を一意に識別する情報であり、当該ノード間通信パケット６００を作成した計算機ノード１００の識別子が設定される。入出力決定処理カウンタ６０２は当該ノード間通信パケット６００を作成した計算機ノード１００の入出力決定処理カウンタ３１５の値である。計算機ノード処理状態６０３は計算機ノード処理状態テーブル７００と同様の構成を有しており、当該ノード間通信パケット６００を作成した計算機ノード１００の処理状態を保持する。計算機ノード処理状態個別詳細情報６０４は特定の構成を持たず、ノード間通信パケット６００を送信する時々の用途に応じた情報が設定される。

＜計算機ノード構成管理テーブル（図７）＞

図７は計算機ノード構成管理テーブル３１７の構成の一例を示す図である

。計算機ノード構成管理テーブル３１７は入力同期処理に寄与した計算機ノード１００の情報を保持するテーブルである。計算機ノード構成管理テーブル３１７は入力同期時計算機ノード総数３１７Ａと計算機ノード識別子６０１、入出力決定処理カウンタ６０２、計算機ノード処理状態６０３の列から構成される。

[0040] 計算機ノード識別子６０１、入出力決定処理カウンタ６０２、計算機ノード処理状態６０３の列には、ノード間通信パケット６００内の同名の列と同様の情報が格納される。計算機ノード構成管理テーブル３１７は、複数計算機ノード分の情報を保持することが可能で、計算機ノード構成管理テーブル３１７に情報が登録されている計算機ノード１００の総数は入力同期時計算機ノード総数３１７Ａに保持される。

＜決定入力管理テーブル（図８）＞

図８は決定入力管理テーブル８００の構成の一例を示す図である。決定入力管理テーブル８００は入力同期処理の結果、アプリケーション２００の処理対象に決定された外部入力パケット５００の情報を保持する。決定入力管理テーブル８００は決定入力総数８０１と、外部入力パケット５００と同様の構成である送信元情報５０１列、シーケンス番号５０２列、送信タイムスタンプ５０３列、順列処理要否フラグ５０４列、順列処理ブロック数５０５列、順列処理開始番号５０６列、コンテンツ５０７内のアプリ識別子５０８列、パラメタ５０９列から構成される。外部入力パケット５００の情報は複数個保持することが可能で、保持数は決定入力総数８０１に保持される。

＜決定済み入力管理テーブル（図９）＞

図９は決定済み入力管理テーブル８１０の構成の一例を示す図である。入力同期処理では、ある外部入力パケット５００を取得した計算機ノード１００が同期システム１内で多数派を占める場合に、当該外部入力パケット５００を取得していない少数派となった計算機ノード１００にも当該外部入力パケット５００情報を共有させて当該計算機ノード１００のアプリケーション２００にも当該外部入力パケットに基づく処理を実行させる。このとき少数

派の計算機ノード１００は、アプリケーション２００実行後に当該外部入力パケット５００を外部ネットワーク３を介して自ら取得する可能性がある。従って、外部入力パケット５００の二重処理を避けるために、処理対象から外したい外部入力パケット５００を決定済み入力管理テーブル８１０に記録する。

[0041] 決定済み入力管理テーブル８１０は決定済み入力総数８１１と、外部入力パケット５００と同様の構成である送信元情報５０１列、シーケンス番号５０２列、送信タイムスタンプ５０３列、順列処理要否フラグ５０４列、順列処理ブロック数５０５列、順列処理開始番号５０６列、コンテンツ５０７内のアプリ識別子５０８列、パラメタ５０９列から構成される。外部入力パケット５００の情報は複数個保持することが可能で、保持数は決定済み入力総数８１１に保持される。

＜決定出力管理テーブル（図１０）＞

図１０は決定出力管理テーブル３２１の構成例を示す図である。決定出力管理テーブル３２１は出力同期処理に寄与した計算機ノード１００の情報を保持するテーブルである。決定出力管理テーブル３２１は出力同期時計算機ノード総数３２１Ａと順位３２１Ｂ、ノード間通信パケット６００内に含まれる列と同一である計算機ノード識別子３２１Ｃ、入出力決定処理カウンタ３２１Ｄ、計算機ノード処理状態３２１Ｅと、計算機ノード処理結果３２１Ｆから構成される。順位３２１Ｂは複数の計算機ノード１００から「リーダー」役になる計算機ノード１００が選択される際の優先順位を意味し、計算機ノード処理結果３２１Ｆはアプリケーション２００の実行結果を保持する。

[0042] 決定出力管理テーブル３２１は複数の計算機ノード１００の情報を保持することが可能であり、保持数は出力同期時計算機ノード総数３２１Ａに保持される。

[0043] 以下、同期システム１が外部入力パケット５００を外部入力パケット管理テーブル５５０に取り込むための処理、及び取り込んだ外部入力パケット５

００に対する一連の処理を図１１を用いて説明する。取り込んだ外部入力パケット５００に対する一連の処理の内、処理対象外部入力パケットの選定を開始する処理を図１２を用いて、処理対象外部入力パケットの選定を行う処理を図１３、図１４、図１５、図１６を用いて、選定した外部入力パケットに対してアプリケーション２００を起動する処理を図１７を用いて、アプリケーション２００の実行結果を決定する処理を図１８、図１９、図２０、図２１、図２２、図２３、図２４を用いて説明する。更に計算機ノード１００の再同期検知処理、再同期処理について図２５、図２６を用いて説明する。なお、各フローチャートの説明ではステップを「Ｓ」と略記する。

＜基盤ソフトウェア処理のフローチャート（図１１）＞

図１１は基盤ソフトウェア処理の一例を示すフローチャートである。基盤ソフトウェア３００は外部入力パケット記憶部３０１と入出力同期部３１０に大別され、それぞれ異なるタスクとしてオペレーティングシステム４００から起動され、両者は独立して動作する。図１１は、外部入力パケット記憶部３０１と入出力同期部３１０、それぞれによって実行される処理のフローチャートを示す。

[0044] 外部入力パケット記憶部３０１は、初めに外部通信I／F１０４による外部入力パケット５００の受信を検出する（Ｓ１０１）。

[0045] 次に外部入力パケット記憶部３０１は受信した外部入力パケット５００の情報を外部入力パケット管理テーブル５５０に格納する（Ｓ１０２）。

[0046] 外部入力パケット保存処理（Ｓ１０２）が終了すると、外部入力パケット記憶部３０１は再び外部入力パケット受信（Ｓ１０１）の実行に移り、前述の処理を繰返す。

[0047] 一方入出力同期部３１０は、初めに再同期要求を行っている他の計算機ノード１００の検出と当該計算機ノード１００についての同期処理を行う。詳細は図２５、図２６と共に後述する（Ｓ１１１）。

[0048] 次に入出力同期部３１０は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の処理状態識別子７０１に「INPUT」値を設定する。同時に

、付加情報７０２には入力同期ステップ７１０としての初期値を設定し、Penalty７０３にも初期値を設定する。入力同期ステップ７１０の初期値には例えば「２」以上の値、一方Penalty７０３の初期値として例えば「０」値が与えられる（Ｓ１１２）。

[0049] 次に入出力同期部３１０は、処理対象となる外部入力パケット５００の選定を開始できるか否かを判別する処理を行う。詳細は図１２と共に後述する（Ｓ１１３）。

[0050] 次に入出力同期部３１０は、処理対象となる外部入力パケット５００を選定する処理を行う。詳細は図１３及び図１４、図１５、図１６と共に後述する（Ｓ１１４）。

[0051] 次に入出力同期部３１０は、アプリケーション２００を実行し、処理対象入力パケット選定処理Ｓ１１４で選定した外部入力パケット５００の処理を行う。詳細は図１７と共に後述する（Ｓ１１５）。

[0052] 次に入出力同期部３１０は、アプリケーション２００による外部入力パケット５００の処理結果を計算機ノード１００間で共有し、端末２へ処理結果を出力する「リーダー」役と出力状態を監視する「モニター」役の計算機ノード１００を決定し、「リーダ」役の計算機ノード１００は端末２へ処理結果を出力する。詳細は図１８及び図１９、図２０、図２１、図２２、図２３、図２４と共に後述する（S１１６）。

[0053] 実行結果決定処理Ｓ１１６が終了すると、入出力同期部３１０は再び再同期検知処理Ｓ１１１を再び実行じ、前述の処理を繰返す。

＜処理対象入力パケット選定開始確認処理のフローチャート（図１２）＞

図１２は、入出力同期部３１０の処理対象入力パケット選定開始確認処理Ｓ１１３の一例を示すフローチャートである。

[0054] 入出力同期部３１０は処理対象入力パケット選定開始確認処理Ｓ１１３を開始すると、自計算機ノード１００に、他の計算機ノード１００から処理対象となる外部入力パケット５００を選定するためのノード間通信パケット６００が到着しているか否かをを確認する（Ｓ２０１及びＳ２０２）。ノード

間通信パケット６００が到着している場合（Ｓ２０２の結果が「Ｙ」）、Ｓ１１４を実行する。ノード間通信パケット６００が到着していない場合（Ｓ２０２の結果が「Ｎ」）、Ｓ２０３を実行する。

[0055] Ｓ２０３及びＳ２０４では、入出力同期部３１０は自計算機ノード１００に到着している外部入力パケット５００の確認を、外部入力パケット管理テーブル５５０を参照することで行う。外部入力パケット５００の存在が確認できる場合（Ｓ２０４の結果が「Ｙ」）、Ｓ２０５で計算機ノード処理状態テーブル７００の処理状態識別子７０１に「INPUT」を設定し、また入力同期ステップ７１０として付加情報７０２に値（例えば２）を設定した後、Ｓ１１４を実行する。

[0056] 外部入力パケット５００の存在が確認できない場合（Ｓ２０４の結果が「Ｎ」）、処理対象となる外部入力パケット５００が存在しないため、入出力同期部３１０は初期実行処理である再同期検知処理Ｓ１１１を実行する。

[0057] 処理対象入力パケット選定処理Ｓ１１４は、図１３、図１４、図１５、図１６に分割して説明する。処理対象入力パケット選定処理Ｓ１１４の開始部分を図１３で説明し、Ｓ１１４に続いて実行される入力決定パケット送信処理Ｓ４００を図１４で説明し、Ｓ４００に続いて実行される入力処理状態一致処理Ｓ５００を図１５で説明し、Ｓ５００に続いて実行される入力多数決処理Ｓ６００を図１６で説明する。

＜処理対象入力パケット選定処理のフローチャート（図１３）＞

図１３は入出力同期部３１０が実行する処理対象入力パケット選定処理Ｓ１１４の一例を示すフローチャートである。

[0058] 入出力同期部３１０は処理対象入力パケット選定処理Ｓ１１４を開始すると、初めに決定入力記憶部３１４の決定入力管理テーブル８００に未処理の外部入力パケット５００の情報が登録されているか否かを確認する（Ｓ３１１）。

[0059] Ｓ３１１で外部入力パケット５００の情報が保持されている場合（Ｓ３１１の結果が「Ｎ」）、当該外部入力パケット５００の情報を優先的に処理す

るために、選定入力パケット処理Ｓ１１５を実行する。

[0060] 外部入力パケットの５００の情報が決定入力管理テーブル８００に保持されていない場合（Ｓ３１１の結果が「Ｙ」）、決定入力記憶部３１４の決定入力管理テーブル８００へ新たな外部入力パケット５００の情報を取り込む処理に移るためにＳ３０１を起動する。

[0061] Ｓ３０１では、入力同期部３１０は処理状態記憶部３１１内の計算機ノード処理状態テーブル７００の処理状態識別子７０１から自計算機ノード１００状態値「INPUT」（Ｓ７０９）を取得する。

[0062] 次に入力同期部３１０は処理状態記憶部３１１内の計算機ノード処理状態テーブル７００の付加情報７０２から入力同期ステップ７１０の値を取得する（Ｓ３０２）。

[0063] 次に入力同期部３１０は入出力決定処理カウンタ３１５の値を取得する（Ｓ３０３）。

[0064] 次に入力同期部３１０はノード間通信パケット６００に値を設定する。計算機ノード識別子６０１には自計算機ノード１００の識別子を設定し、入出力決定処理カウンタ６０２にはＳ３０３で取得した入出力決定処理カウンタ３１５を設定し、計算機ノード処理状態６０３にはＳ３０１とＳ３０２で取得した計算機ノード処理状態テーブル７００の計算機ノード状態情報を設定する（Ｓ３０４）。

[0065] 次に入力同期部３１０は外部入力パケット記憶部３０１の外部入力パケット管理テーブル５５０から入力同期ステップ７１０数分の外部入力パケット５００の情報を取得する（Ｓ３０５）。

[0066] Ｓ３０５で入力同期部３１０は、取得した外部入力パケット５００の情報の中に決定済み入力管理テーブル８１０内に保持されている外部入力パケット５００の情報と一致するものが存在しないことを確認する（Ｓ３０６およびＳ３０７）。

[0067] Ｓ３０５で入力同期部３１０は、取得した外部入力パケット５００の情報と同じ情報が決定済み入力管理テーブル８１０内に存在する場合（Ｓ３０７

の結果が「Ｎ」）、外部入力パケット管理テーブル５５０から当該外部入力パケット５００の情報を削除し（Ｓ３０８）、削除した分と同数の外部入力パケット５０００の情報を外部入力パケット管理テーブル５５０から取り出す（Ｓ３０５）。

[0068] Ｓ３０５で入力同期部３１０が取得した外部入力パケット５００の情報と同じ情報が決定済み入力管理テーブル８１０内に存在しない場合（Ｓ３０７の結果が「Ｙ」）、Ｓ３０９を実行する。

[0069] Ｓ３０９で入力同期部３１０は、取得した外部入力パケット５００の情報をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定する。以上でノード間通信パケット６００の作成が完了すると、入力決定パケット送信処理Ｓ４００を実行する。

＜入力決定パケット送信処理のフローチャート（図１４）＞

図１４は、入出力同期部３１０の処理対象入力パケット選定処理Ｓ１１４の一部分である、入力決定パケット送信処理Ｓ４００の一例を示すフローチャートである。

[0070] 入出力同期部３１０は入力決定パケット送信処理Ｓ４００を開始すると、初めに処理状態記憶部３１１内の計算機ノード処理状態テーブル７００の処理状態７０４の値を「BEGIN」に設定する（同時に自計算機ノード１００状態をコピーしたノード間通信パケット６００の計算機ノード処理状態６０３の値も更新する）（Ｓ４０１）。

[0071] 次に入出力同期部３１０は、作成したノード間通信パケット６００をノード間通信Ｉ／Ｆ１０５を介して他の計算機ノード１００にマルチキャストで送信する（Ｓ４０２）。

[0072] 次に入出力同期部３１０は、他の計算機ノード１００からノード間通信パケット６００が到着することを期待して、タイムアウト時間までＷａｉｔする（Ｓ４０３）。

[0073] 次に入出力同期部３１０は、他の計算機ノード１００から送信されたノード間通信パケット６００を受信する（Ｓ４０４）。

[0074] 次に入出力同期部３１０は、自計算機ノード１００で作成したノード間通信パケット６００情報と、Ｓ４０４で受信した他計算機ノード１００のノード間通信パケット６００情報を結合し、「受信結果リスト」を作成する。「受信結果リスト」はノード間通信パケット６００と同様の列情報を持つ（Ｓ４０５）。各計算機ノードの結果を結合処理することで、一部の内部通信経路の障害により通信不可能な計算機ノードが存在した場合でも通信可能な経路より全計算機ノードの情報が共有可能となる。

[0075] 次に入出力同期部３１０は、Ｓ４０５で作成した受信結果リスト中の、計算機ノード識別子６０１、入出力決定処理カウンタ６０２、計算機ノード処理状態６０３の列情報を計算機ノード構成管理テーブル３１７の同じ列にコピーする。この時、入力同期時計算機ノード総数３１７Ａを更新する。（Ｓ４０６）

以上で入力決定パケット送信処理Ｓ４００が完了すると、入力処理状態一致処理Ｓ５００を実行する。

＜入力処理状態一致処理のフローチャート（図１５）＞

図１５は、入出力同期部３１０の処理対象入力パケット選定処理Ｓ１１４の一部分である、入力処理状態一致処理Ｓ５００の一例を示すフローチャートである。

[0076] 入力処理状態一致処理Ｓ５００では、各計算機ノード１００間で「入力同期」処理における状態を一致させる処理を行う。同期システム１は各計算機ノード１００の演算結果を多数決判定することで端末２への出力情報の信頼性を確保するが、各計算機ノード１００は独立した計算機であるため、各自が異なる計算機ノード状態で稼動している。演算結果の多数決判定を実施するためには、複数の各計算機ノード１００の演算結果を必要とする。演算結果の多数決判定に参加する計算機ノード１００の数を最大化するために、演算処理の入力段階で一度計算機ノード１００間の同期を取る。入力処理状態一致処理Ｓ５００の方針は、処理周期が明らかに異なる計算機ノード１００は排除するが、僅かな計算機ノード処理状態の遅れはこれを修正することで

演算処理、演算結果の多数決判定に参加する計算機ノード１００数を最大化する。

[0077] 入出力同期部３１０は入力処理状態一致処理Ｓ５００を開始すると、初めに入出力決定処理カウンタ３１５と、Ｓ４０４で他の計算機ノード１００から受信したノード間通信パケット６００の入出力決定処理カウンタ６０２を比較し、自他計算機ノード１００の入出力決定処理カウンタ３１５の値が一致しているかを判定する。計算機ノード１００は複数の他の計算機ノード１００と送受信を行っている場合があるため、受信したノード間通信パケット６００も複数個になることがある。Ｓ５０２での「比較」とは、入出力決定処理カウンタ３１５値と受信した一又は複数の入出力決定処理カウンタ６０２値に基づいて自他計算機ノード１００を分類し、自計算機ノード１００が多数派または少数派のいずれに分類されるかを決定する、「多数決判定」のことである（Ｓ５０１およびＳ５０２）。

[0078] 尚「多数決判定」において、分類の結果３つ以上のグループが生成された場合には、最も多くの計算機ノードが属するグループに自計算機ノードが分類された場合に、自計算機ノードが「多数派」に分類されたと判定してもよい。これは、以降の処理において実行される様々な「多数決判定」においても同様である。

[0079] 自計算機ノード１００が少数派に分類される場合（Ｓ５０２の結果が「Ｎ」）、多数派の計算機ノード１００と処理周期が異なるため、再同期処理Ｓ１１８を実行する。

[0080] 自計算機ノード１００が多数派に分類される場合（Ｓ５０２の結果が「Ｙ」）、処理周期が一致しているとして、継続するＳ５０４を実行する。

[0081] 次に自計算機ノード１００が他の計算機ノード１００と計算機ノード処理状態が一致しているかを比較する。ここでの「比較」も「多数決判定」のことである。計算機ノード処理状態の比較は、他計算機ノード１００から受信したノード間通信パケット６００の計算機ノード処理状態６０３（計算機ノード処理状態６０３の構成列は計算機ノード処理状態テーブル７００と同じ

）と自計算機ノード１００の計算機ノード処理状態テーブル７００の値（処理状態識別子７０１、処理状態７０４）を比較することで行う（Ｓ５０４およびＳ５０５）。

[0082] Ｓ５０５で自計算機ノード１００が少数派に分類される場合（Ｓ５０５の結果が「Ｎ」）、計算機ノード処理状態が多数派の計算機ノード処理状態と一致していないため、Ｓ５０６を実行する。Ｓ５０６では自計算機ノード１００の計算機ノード処理状態を多数派の計算機ノード処理状態と合わせるために、計算機ノード処理状態テーブル７００（処理状態識別子７０１、付加情報７０２、処理状態７０４）を多数派の計算機ノード処理状態値（計算機ノード処理状態６０３）で更新し、また計算機ノード処理状態テーブル７００のPenalty７０３を設定（例えば「１」を設定）する。

[0083] 次に自計算機ノード処理状態テーブル７００の処理状態７０４に値が、「BEGIN」の場合（Ｓ５０７の結果が「Ｙ」）はＳ５０８を実行し、「END」の場合（Ｓ５０７の結果が「Ｎ」）は入力処理状態一致処理Ｓ５００を終了して入力多数決処理Ｓ６００を実行する（Ｓ５０６およびＳ５０７）。

[0084] Ｓ５０５で自計算機ノード１００が多数派に分類される場合（Ｓ５０５の結果が「Ｙ」）、Ｓ５０８を実行する。

[0085] Ｓ５０８では、自計算機ノード処理状態テーブル７００の処理状態７０４を「END」に変更する。各計算機ノード１００は、それぞれの計算機ノード１００の入出力同期部３１０が処理候補として取得した外部入力パケット５００の情報（受信結果リスト）の共有が完了した状態となる（Ｓ５０８）。

[0086] 次にノード間通信パケット６００に設定した計算機ノード処理状態テーブル７００値、及び「受信結果リスト」を、ノード間通信Ｉ／Ｆ１０５経由で他計算機ノード１００にマルチキャストで送信する（Ｓ５０９）。

[0087] 次に他計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ５１０）。

[0088] Ｗａｉｔ終了で、入力処理状態一致処理Ｓ５００を完了し、次の入力多数決処理Ｓ６００を実行する。

＜入力多数決処理のフローチャート（図１６）＞

図１６は入出力同期部３１０の処理対象入力パケット選定処理Ｓ１１４の一部分である、入力多数決処理Ｓ６００の一例を示すフローチャートである。入力多数決処理Ｓ６００ではアプリケーション２００の処理対象となる外部入力パケット５００を決定する。

[0089] 入出力同期部３１０は入力多数決処理Ｓ６００を開始すると、初めに「受信結果リスト」を参照して、過半数の計算機ノード１００に共通して取得されている外部入力パケット５００（「多数決同意入力パケット」と称する）の有無を確認する（Ｓ６０１）。ここで過半数とは、現時点で計算機ノード構成管理テーブルに情報が登録されている計算機ノードの数の過半数であっても良いし、同期システム１に存在する計算機ノードの数の過半数であっても良い。

[0090] 多数決同意入力パケットが存在しない場合（Ｓ６０１の結果が「Ｎ」）、Ｓ６０２を実行して処理状態記憶部３１１の計算機ノード処理状態テーブル７００の付加情報７０２（入力同期ステップ７１０）の値を「１」増加し、処理対象入力パケット選定処理Ｓ１１４の開始位置（Ｓ３００）に実行を移す。

[0091] 多数決同意入力パケットが存在する場合（Ｓ６０１の結果が「Ｙ」）、次処理Ｓ６０４を実行する。

[0092] 次にＳ６０１で選定された外部入力パケット５００情報の順序の並び替えを行う。並び替えた外部入力パケット５００を「決定入力パケット」と称する。並び替えは外部入力パケット５００のシーケンス番号５０２の値を基に行う（Ｓ６０４）。

[0093] 次にＳ６０４で並び替えた外部入力パケット５００の情報を決定入力記憶部３１４の決定入力管理テーブル８００に保存し、決定入力総数８０１を更新する（Ｓ６０５）。

[0094] 次にＳ６０５で決定入力管理テーブル８００に保存された各外部入力パケット５００の情報について、過半数の計算機ノード１００が図１３のＳ３０

５で取得した外部入力パケット５００情報と同一の情報を自計算機ノード１００が取得を行ったかどうかの多数決判定を行う（Ｓ６０６）。

[0095] 自計算機ノード１００が同一の情報を取得していない場合（Ｓ６０６の結果が「Ｎ」）、Ｓ６０７で当該外部入力パケット５００の情報を決定入力記憶部３１４の決定済み入力管理テーブル８１０に設定し、決定済み入力総数８１１を更新して、Ｓ６０９を起動する。当該外部入力パケット５００の情報はアプリケーション２００の処理対象であるが、自計算機ノード１００が自ら外部ネットワーク３から取得した情報では無い為、後の演算周期で遅れて到着した当該外部入力パケット５００を外部ネットワーク３から取得する可能性がある。同一外部入力パケット５００に対する多重処理を回避するため、当該外部入力パケット５００の情報を決定済み入力管理テーブル８１０に設定しておく。

[0096] 自計算機ノード１００が同一の情報取得している場合（Ｓ６０６の結果が「Ｙ」）、Ｓ６０６を起動して「決定入力パケット」のうち、自計算機ノード１００が自ら外部ネットワーク３経由で取得した外部入力パケット５００の情報を、外部入力パケット記憶部３０１の外部入力パケット管理テーブル５５０から削除し、続いてＳ６０９を起動する。

[0097] Ｓ６０９は入出力決定処理カウンタ３１５を「１」増加する。

[0098] 次に決定入力パケットをノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定する（Ｓ６１０）。

[0099] 次にＳ６１０で作成したノード間通信パケット６００を、ノード間通信Ｉ／Ｆ１０５経由で他の計算機ノード１００にマルチキャストで送信する（Ｓ６１１）。

[0100] 次に他計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ６１２）。

[0101] Ｗａｉｔ終了で、入力多数決処理Ｓ６００を完了し、次の選定入力パケット処理Ｓ１１５を実行する。

＜選定入力パケット処理のフローチャート（図１７）＞

図１７は、入出力同期部３１０の処理対象入力パケット選定処理Ｓ１１４の一部分である、選定入力パケット処理Ｓ１１５の一例を示すフローチャートである。選定入力パケット処理Ｓ１１５ではアプリケーション２００を起動し、「決定入力パケット」を処理する。

[0102] 入出力同期部３１０は選定入力パケット処理Ｓ１１５を開始すると、初めに決定入力記憶部３１４の決定入力管理テーブル８００から外部入力パケット５００情報を取り出す（Ｓ７０１）。

[0103] Ｓ７０１で取得した外部入力パケット５００の情報うち、コンテンツ５０７のアプリ識別子５０８から、処理を実行すべきアプリケーション２００を決定し、パラメタ５０９を付加して該当アプリケーション２００の開始処理（Ｓ７０２－Ａ）を起動する。アプリケーション２００の起動後、選定入力パケット処理Ｓ１１５は当該アプリケーション２００の処理終了待ちとなる。（Ｓ７０２）

Ｓ７０２で起動されたアプリケーション２００は付加パラメタ５０９を取得して処理を開始し（Ｓ７０２－Ａ）、アプリケーション２００処理を実行し（Ｓ７０２－Ｂ）、終了時には処理結果を選定入力パケット処理Ｓ１１５に出力する（Ｓ７０２－Ｃ）。

[0104] アプリケーション２００処理終了後、選定入力パケット処理Ｓ１１５は処理結果を取得する（Ｓ７０３）。

[0105] 外部入力パケット５００の情報に応じたアプリケーション２００の起動と処理結果の取得で選定入力パケット処理Ｓ１１５を完了し、次の実行結果決定処理Ｓ１１６を実行する。

[0106] 実行結果決定処理Ｓ１１６は、本書では図１８、図１９、図２０、図２１、図２２、図２３、図２４に分割して説明する。実行結果決定処理Ｓ１１６の開始部分を図１８で説明し、Ｓ１１６に続いて実行される出力処理状態一致処理Ｓ９００を図１９で説明し、Ｓ９００に続いて実行される出力多数決処理Ｓ１０００を図２０で説明し、Ｓ１０００に続いて実行されるリーダー選出処理Ｓ１１００を図２１で説明し、Ｓ１１００に続いて実行されるリー

ダー処理状態一致化処理Ｓ１２００を図２２で説明し、Ｓ１２００に続いて実行されるボーター処理Ｓ１３００を図２３で説明し、Ｓ１３００に続いて実行される出力異常処理Ｓ１４００を図２４で説明する。

＜実行結果決定処理のフローチャート（図１８）＞

図１８は入出力同期部３１０の実行結果決定処理Ｓ１１６の一例を示すフローチャートである。

[0107] 入出力同期部３１０は実行結果決定処理Ｓ１１６を開始すると、初めに処理状態記憶部３１１の計算機ノード処理状態テーブル７００の処理状態識別子７０１を「OUTPUT」に、付加情報７０２を「OUTPUT」に、処理状態７０４を「BEGIN」に更新する（Ｓ８０１）。

[0108] 次に入出力同期部３１０は、入出力決定処理カウンタ３１５の値を確認する（Ｓ８０２）。

[0109] 次に入出力同期部３１０は、Ｓ７０３で取得したアプリケーション２００の処理結果を、ノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定する。この時、Ｓ８０１で更新した計算機ノード処理状態テーブル７００値を計算機ノード処理状態６０３に設定し、Ｓ８０２で取得した入出力決定処理カウンタ３１５値を入出力決定処理カウンタ６０２にも設定する。（Ｓ８０３）

次に入出力同期部３１０は、Ｓ８０３で作成したノード間通信パケット６００をノード間通信Ｉ／Ｆ１０５経由で他の計算機ノード１００にマルチキャストで送信することで、アプリケーション２００の処理結果を他の計算機ノード１００との間で共有する（Ｓ８０４）。

[0110] 次に入出力同期部３１０は、他の計算機ノード１００からのノード間通信パケット６００が到着を待つために、タイムアウトまでＷａｉｔする（Ｓ８０５）。

[0111] Ｓ８０６で入出力同期部３１０は、他の計算機ノード１００から到着したノード間通信パケット６００を取得し、Ｓ８０７で自他計算機ノード１００のノード間通信パケット６００を結合して「受信結果リスト」を作成する。

[0112] 更に入出力同期部３１０は、受信結果リスト（構成列はノード間通信パケット６００に同じ）に登録されている計算機ノード識別子６０１、入出力決定処理カウンタ６０２、計算機ノード処理状態６０３の列情報を計算機ノード構成管理テーブル３１７に設定する。ただし入力同期時計算機ノード総数３１７Ａの値は更新しない。（Ｓ８０８）

計算機ノード１００間でのアプリケーション２００の処理結果の交換を行ったところで、実行結果決定処理Ｓ１１６を完了し、次の出力処理状態一致処理Ｓ９００を実行する。

＜出力処理状態一致処理のフローチャート（図１９）＞

図１９は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、出力処理状態一致処理Ｓ９００の一例を示すフローチャートである。

[0113] 出力処理状態一致処理Ｓ９００では、各計算機ノード１００間で「出力同期」処理における状態を一致させる処理を行う。同期システム１は各計算機ノード１００の演算結果を多数決判定することで端末２への出力情報の信頼性を確保するが、各計算機ノード１００は独立した計算機であるため、各自が異なる計算機ノード状態で稼動している。出力処理状態一致処理Ｓ９００の方針は、処理周期が明らかに異なる計算機ノード１００は排除するが、僅かな計算機ノード処理状態の遅れはこれを修正することで演算結果の多数決判定に参加する計算機ノード１００数を最大化する。

[0114] 入出力同期部３１０は出力処理状態一致処理Ｓ９００を開始すると、初めに入出力決定処理カウンタ３１５と、Ｓ８０６で他の計算機ノード１００から受信したノード間通信パケット６００の入出力決定処理カウンタ６０２を比較し、自他計算機ノード１００の入出力決定処理カウンタ３１５の値が一致しているかを判定する。計算機ノード１００は複数の他の計算機ノード１００とノード間通信パケット６００を送受信する場合があるため、受信したノード間通信パケット６００も複数個になることがある。Ｓ９０１での「比較」とは、入出力決定処理カウンタ３１５値と複数の入出力決定処理カウンタ６０２値とに基づき、自他計算機ノード１００を分類し、自計算機ノード

１００が多数派または少数派のいずれに分類されるかを決定する、「多数決判定」のことである（Ｓ９０１およびＳ９０２）。

[0115] 自計算機ノード１００が少数派に分類される場合（Ｓ９０２の結果が「Ｎ」）、多数派の計算機ノード１００と処理周期が異なるため、入出力同期部３１０は再同期処理Ｓ１１８を実行する。

[0116] 自計算機ノード１００が多数派に分類される場合（Ｓ９０２の結果が「Ｙ」）、処理周期が一致しているとして、入出力同期部３１０は継続するＳ９０４を実行する。

[0117] 次に自計算機ノード１００が他の計算機ノード１００と計算機ノード処理状態が一致しているかを比較する。ここでの「比較」も「多数決判定」のことである。計算機ノード処理状態の比較は、他計算機ノード１００から受信したノード間通信パケット６００の計算機ノード処理状態６０３（構成列は計算機ノード処理状態テーブル７００と同じ）と自計算機ノード１００の計算機ノード処理状態テーブル７００の値（処理状態識別子７０１、付加情報７０２、処理状態７０４）を比較することで行う。（Ｓ９０４およびＳ９０５）

Ｓ９０５で自計算機ノード１００が少数派に分類される場合（Ｓ９０５の結果が「Ｎ」）、計算機ノード処理状態が多数派の計算機ノード処理状態と一致していないため、入出力同期部３１０はＳ９０６を実行する。

[0118] Ｓ９０６では、他計算機ノード１００から受信したノード間通信パケット６００の計算機ノード処理状態６０３（構成列は計算機ノード処理状態テーブル７００と同じ）の付加情報７０２を参照し、多数派の計算機ノード１００の計算機ノード処理状態が「LEADER」にあるかを判別する。多数派の計算機ノード処理状態が「LEADER」にある場合（Ｓ９０６の結果が「Ｙ」）、「出力同期」中の一つ先の処理で行われる、端末２へ処理結果を出力する候補の計算機ノード１００が決定済みであるため、自計算機ノード１００は監視役として動作するために入出力同期部３１０はボーター処理Ｓ１３００を起動する。多数派の計算機ノード処理状態が「LEADER」にない場合（Ｓ９０６

の結果が「Ｎ」）、入出力同期部３１０はＳ９０７を起動する。また、多数派の計算機ノード１００の付加情報７０２の値の如何に関わらず、入出力同期部３１０は自計算機ノード１００の計算機ノード処理状態テーブル７００のPenalty７０３を設定（例えば「１」を設定）する。

[0119] Ｓ９０６の判定において、多数派の計算機ノード１００の計算機ノード処理状態が「VOTER」であるケースもあり得る。しかしS９０６においては「LEADER」のみを検出対照とする。多数派の計算機ノード１００の計算機ノード処理状態が「VOTER」である場合、これら多数派の計算機ノード１００は自計算機ノード処理状態「OUTPUT」よりも「出力同期」処理において二つ先の、端末２に結果出力を行う処理状態にある。多数派計算機ノード１００の処理状態が「VOTER」であることを確認した後に、自計算機ノード１００の処理状態を「VOTER」に補正して出力監視役である「モニター」に切替えたとしても、端末２への結果出力を見逃す可能性があることから、Ｓ９０６の判定は「LEADER」のみを検出対象とする。つまり、「出力同期」の処理段階における「OUTPUT」から「VOTER」への二段階の同期ずれの補正は行わない。

[0120] Ｓ９０７とＳ９０８では多数派の計算機ノード処理状態の比較を行い、付加情報７０２が「OUTPUT」あり（Ｓ９０７の結果が「Ｙ」）、更に処理状態７０４が「END」である場合（Ｓ９０８の結果が「Ｙ」）は計算機ノード１００間でアプリケーション２００の処理結果を一致化させる処理が完了しているため、入出力同期部３１０は出力多数決処理Ｓ１０００を起動して出力処理状態一致処理Ｓ９００を完了する。

[0121] 一方付加情報７０２が「OUTPUT」でない場合（Ｓ９０７の結果が「Ｎ」）、または付加情報７０２は「OUTPUT」ある（Ｓ９０７の結果が「Ｙ」）が処理状態７０４が「END」でない場合（Ｓ９０８の結果が「Ｎ」）は、入出力同期部３１０はタイムアウトまでＷａｉｔした後（Ｓ９０９）、Ｓ８０６の他計算機ノード１００からのアプリケーション２００処理結果受信処理から、出力処理状態一致処理Ｓ９００を再実行する。

[0122] 一方、Ｓ９０５で自計算機ノード１００が多数派に分類される場合（Ｓ９

０５の結果が「Ｙ」）、入出力同期部３１０は次のＳ９１１を起動する。

[0123] Ｓ９１１では自計算機ノード処理状態テーブル７００の処理状態７０４を「END」に変更する。各計算機ノード１００は、それぞれの計算機ノード１００のアプリケーション２００の処理結果の共有が完了した状態となる。

[0124] 次にノード間通信パケット６００に設定した計算機ノード処理状態テーブル７００値をノード間通信Ｉ／Ｆ１０５経由で他計算機ノード１００にマルチキャストで送信する（Ｓ９１２）。

[0125] 次に他計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ９１３）。

[0126] Ｗａｉｔ終了で、出力処理状態一致処理Ｓ９００を完了し、次の出力多数決処理Ｓ１０００を実行する。

＜出力多数決処理のフローチャート（図２０）＞

図２０は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、出力多数決処理Ｓ１０００の一例を示すフローチャートである。出力多数決処理Ｓ１０００では、各計算機ノード１００が算出したアプリケーション２００の処理結果の多数決判定を行う。

[0127] 入出力同期部３１０は出力多数決処理Ｓ１０００を開始すると、初めに「受信結果リスト」中の各計算機ノード１００が算出したアプリケーション２００の処理結果の中に、過半数の計算機ノード１００に共通して存在する処理結果（「多数決同意出力」と称する）が存在するか、その有無を判定する（Ｓ１００１）。ここで過半数とは、計算機ノード構成テーブル３１７に現時点で登録されている計算機ノードの数の過半数であっても良いし、同期システム１に存在する計算機ノードの数の過半数であっても良い。

[0128] 多数決同意出力が存在しない場合（Ｓ１００１の結果が「Ｎ」）、アプリケーション２００の処理結果の多数決判定が不能ということになり、入出力同期部３１０はＳ１００２を起動する。

[0129] Ｓ１００２は「受信結果リスト」中の計算機ノード１００総数と、計算機ノード構成管理テーブル３１７の入力同期時計算機ノード総数３１７Ａを比

較する。総数が一致する場合（Ｓ１００２の結果が「Ｙ」）、各計算機ノード１００が算出したアプリケーション２００の処理結果に二つとして同じものが無かったという事になり、入出力同期部３１０はＳ１００３を起動して基盤ソフトウェア３００の処理を終了させる。総数が一致しない場合（Ｓ１００２の結果が「Ｎ」）、まだアプリケーション２００の処理結果を受信していない計算機ノード１００が存在するため、入出力同期部３１０はＳ８０６の他計算機ノード１００からのアプリケーション２００処理結果の受信処理から、出力処理状態一致処理Ｓ９００を再開する。

[0130] 一方、多数決同意出力が存在する場合（Ｓ１００１の結果が「Ｙ」）、入出力同期部３１０はＳ１００５を起動して自計算機ノード１００のアプリケーション２００の処理結果が多数派に属するかを判定する。自計算機ノード１００が多数派に属する場合（Ｓ１００５の結果が「Ｙ」）、入出力同期部３１０は出力多数決処理Ｓ１０００を完了し、次のリーダー選出処理Ｓ１１００を実行する。自計算機ノード１００が少数派に属する場合（Ｓ１００５の結果が「Ｎ」）、自計算機ノード１００は多数派の計算機ノード１００とは異なる処理結果を導いたという事になり、端末２への結果出力を行う「リーダー」にも結果出力の監視を行う「モニター」にもなり得ないため入出力同期部３１０は今周期の処理からは外れ、次周期に向けて再同期処理Ｓ１１８を起動する。

＜リーダー選出処理のフローチャート（図２１）＞

図２１は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、リーダー選出処理Ｓ１１００の一例を示すフローチャートである。リーダー選出処理Ｓ１１００は、端末２への結果出力を行う「リーダー」候補となる計算機ノード１００の選出を行う。

[0131] 入出力同期部３１０はリーダー選出処理Ｓ１１００を開始すると、初めに処理状態記憶部３１１の計算機ノード処理状態テーブル７００の付加情報７０２を「LEADER」に、処理状態７０４を「BEGIN」に変更する（Ｓ１１０１）。

[0132] 次に入出力同期部３１０は、受信結果リストより「リーダー」候補となる計算機ノード１００を選出し、順位を決定し、「リーダー選出結果」を作成する。選出及び順位の決定方法は、一例として、計算機ノード処理状態テーブル７００のPenalty７０３の値が０でありノード間通信パケット６００の計算機ノード識別子６０１が最も大きな値を持つ順に「リーダー」候補を決定する方法があげられる。（Ｓ１１０２）

次に入出力同期部３１０は、入出力決定処理カウンタ３１５を取得しノード間通信パケット６００の入出力決定処理カウンタ６０２に設定する（Ｓ１１０３）。またＳ１１０２の「リーダー選出結果」をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定する（Ｓ１１０４）。

[0133] 次に入出力同期部３１０は、Ｓ１１０４で作成したノード間通信パケット６００をノード間通信Ｉ／Ｆ１０５経由で他の計算機ノード１００にマルチキャストで送信する（Ｓ１１０５）。

[0134] 次に入出力同期部３１０は、他計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ１１０６）。

[0135] Ｓ１１０７で他計算機ノード１００からのノード間通信パケット６００を取得すると、入出力同期部３１０はＳ１１０８で自他計算機ノード１００のノード間通信パケット６００を結合し、「受信結果リスト」を作成する。「受信結果リスト」は自他計算機ノード１００が決定した「リーダー選出結果」を含む。

[0136] Ｓ１１０９で入出力同期部３１０は、計算機ノード構成管理テーブル３１７に「受信結果リスト」内容を反映して、リーダー選出処理Ｓ１１００を完了し、次のリーダー処理状態一致化処理Ｓ１２００を実行する。

＜リーダー処理状態一致化処理のフローチャート（図２２）＞

図２２は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、リーダー処理状態一致化処理Ｓ１２００の一例を示すフローチャートであ

る。リーダー処理状態一致化処理Ｓ１２００は、計算機ノード１００間でリーダー選出処理の完了状態を一致させる処理を行う。

[0137] 入出力同期部３１０はリーダー処理状態一致化処理Ｓ１２００を開始すると、初めに入出力決定処理カウンタ３１５と、Ｓ１１０７で他の計算機ノード１００から受信したノード間通信パケット６００の入出力決定処理カウンタ６０２を比較し、自他計算機ノード１００の入出力決定処理カウンタ３１５の値が一致しているかを判定する。受信したノード間通信パケット６００は複数個になることがあるため、Ｓ１２０２は「多数決判定」を行う（Ｓ１２０１およびＳ１２０２）。

[0138] 自計算機ノード１００が少数派に分類される場合（Ｓ１２０２の結果が「Ｎ」）、多数派の計算機ノード１００と処理周期が異なるため、入出力同期部３１０は再同期処理Ｓ１１８を実行する。

[0139] 自計算機ノード１００が多数派に分類される場合（Ｓ１２０２の結果が「Ｙ」）、処理周期が一致しているとして、入出力同期部３１０は継続するＳ１２０４を実行する。

[0140] 次に入出力同期部３１０は入出力同期部３１０は自計算機ノード１００が他の計算機ノード１００と計算機ノード処理状態が一致しているかを比較する。ここでの「比較」も「多数決判定」のことである。計算機ノード処理状態の比較は、他計算機ノード１００から受信したノード間通信パケット６００の計算機ノード処理状態６０３（構成列は計算機ノード処理状態テーブル７００と同じ）と自計算機ノード１００の計算機ノード処理状態テーブル７００の値（処理状態識別子７０１、付加情報７０２、処理状態７０４）を比較することで行う（Ｓ１２０４およびＳ１２０５）。

[0141] 自計算機ノード１００が少数派に分類される場合（Ｓ１２０５の結果が「Ｎ」）、他の計算機ノード１００と共にアプリケーション２００の処理結果を端末２に出力し得る状態に遷移出来ないため、入出力同期部３１０は処理周期を離れて再同期処理Ｓ１１８を実行する。

[0142] 自計算機ノード１００が多数派に分類される場合（Ｓ１２０５の結果が「

Ｙ」）、入出力同期部３１０は次のＳ１２０６を実行する。

[0143] 次に入出力同期部３１０は自計算機ノード処理状態テーブル７００の処理状態７０４を「END」に変更する。各計算機ノード１００は、それぞれの計算機ノード１００の入出力同期部３１０が「リーダー選出結果」の共有が完了した状態となる（Ｓ１２０６）。

[0144] 次に入出力同期部３１０は、計算機ノード処理状態テーブル７００値及び「受信結果リスト」を設定したノード間通信パケット６００を、ノード間通信Ｉ／Ｆ１０５経由で他計算機ノード１００にマルチキャストで送信する（Ｓ１２０７）。

[0145] 次に入出力同期部３１０は他計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ１２０８）。

[0146] Ｗａｉｔ終了で、入出力同期部３１０はリーダー処理状態一致化処理Ｓ１２００を完了し、次のボーター処理Ｓ１３００を実行する。

＜ボーター処理のフローチャート（図２３）＞

図２３は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、ボーター処理Ｓ１３００の一例を示すフローチャートである。

[0147] ボーター処理Ｓ１３００では、リーダー選出処理Ｓ１１００で決定した「リーダー選出結果」に従い「リーダー」役と「モニター」役の計算機ノード１００を決定する。「リーダー」役の計算機ノード１００は出力多数決決定処理Ｓ１０００で得た「多数決同意出力」を端末２に出力し、出力状態を「モニター」役の計算機ノード１００が監視する。出力状態に異常が認められた場合、「モニター」役の計算機ノード１００の中から新たな「リーダー」計算機ノード１００を決定し、「多数決同意出力」の端末２への出力を再実行する（出力異常処理Ｓ１４００）。出力異常処理Ｓ１４００の詳細は図２４と共に後述する。

[0148] 入出力同期部３１０はボーター処理Ｓ１３００を開始すると、初めにＳ１２０７で送受信した「リーダー選出結果」の「受信結果リスト」内容で、決

定出力管理テーブル３２１の構成列（出力同期時計算機ノード総数３２１Ａ、順位３２１Ｂ、計算機ノード識別子３２１Ｃ、入出力決定処理カウンタ３２１Ｄ、計算機ノード処理状態３２１Ｅ）及び出力同期時計算機ノード総数３２１Ａの値を更新する（Ｓ１３０１）。

[0149] 次に入出力同期部３１０は、決定出力管理テーブル３２１の順位３２１Ｂ及び計算機ノード識別子３２１Ｃの値より、自計算機ノード１００が「リーダー」であるかを判定する（Ｓ１３０２）。

[0150] 自計算機ノード１００が「リーダー」であると判定される場合（Ｓ１３０２の結果が「Ｙ」）、「リーダー」として動作するために、入出力同期部３１０はＳ１３０４を実行する。自計算機ノード１００が「リーダー」であると判定されない場合（Ｓ１３０２の結果が「Ｎ」）、「モニター」として動作するために、入出力同期部３１０はＳ１３０３を実行する。

[0151] 「リーダー」計算機ノード１００は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の付加情報７０２を「VOTER」に更新し、また処理状態７０４を「LEADER」に更新することで、「リーダー」動作を開始する（Ｓ１３０４）。そして入出力決定処理カウンタ３１５の値を「１」増加させる（Ｓ１３０５）。

[0152] 次にノード間通信パケット６００の入出力決定処理カウンタ６０２にＳ１３０５で更新した入出力決定処理カウンタ３１５値を設定し、計算機ノード処理状態６０３にはＳ１３０４で更新した計算機ノード処理状態テーブル７００値を設定する（Ｓ１３０６）。

[0153] 次にノード間通信パケット６００をノード間通信Ｉ／Ｆ１０５経由で他の計算機ノード１００にマルチキャストで送信し、自計算機ノード１００が「リーダー」であることを通知する（Ｓ１３０７）。そして他の計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ１３０８）。

[0154] Ｗａｉｔ終了後、アプリケーション２００の処理結果を外部通信I／Ｆ１０４経由で端末２にマルチキャストで送信する。マルチキャストによる送信で

あるため、アプリケーション２００の処理結果は、「モニター」計算機ノード１００の外部通信I／Ｆ１０４にも到達する（Ｓ１３０９）。

[0155] 「リーダー」計算機ノード１００は、端末２に送信することで「リーダー」動作を終了し、以上でボーター処理Ｓ１３００を完了して次の再同期検知処理Ｓ１１１を実行する。

[0156] 「モニター」計算機ノード１００は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の付加情報７０２を「VOTER」に更新して、また処理状態７０４を「MONITOR」に更新することで、「モニター」動作を開始する（Ｓ１３１１）。そして入出力決定処理カウンタ３１５の値を「１」増加させる（Ｓ１３１２）。

[0157] 次にノード間通信パケット６００の入出力決定処理カウンタ６０２にＳ１３１２で更新した入出力決定処理カウンタ３１５値を設定し、計算機ノード処理状態６０３にはＳ１３１１で更新した計算機ノード処理状態テーブル７００値を設定する（Ｓ１３１３）。

[0158] 次にノード間通信パケット６００をノード間通信Ｉ／Ｆ１０５経由で他の計算機ノード１００にマルチキャストで送信し、自計算機ノード１００が「モニター」であることを通知する（Ｓ１３１４）。そして他の計算機ノード１００からのノード間通信パケット６００の到着を待つために、タイムアウトまでＷａｉｔする（Ｓ１３１５）。

[0159] Ｗａｉｔ終了後、「リーダー」計算機ノード１００によるアプリケーション２００の処理結果送信状態を、「モニター」計算機ノード１００の外部通信I／Ｆ１０４経由で監視し、異常が発生していないかを判別する（Ｓ１３１６およびＳ１３１７）。

[0160] 「リーダー」計算機ノード１００の処理結果送信に異常が確認できる場合（Ｓ１３１７の結果が「Ｙ」）、出力異常処理Ｓ１４００を起動し、端末２へのアプリケーション２００の処理結果送信をリトライする。

[0161] 「リーダー」計算機ノード１００の処理結果送信に異常が確認できない場合（Ｓ１３１７の結果が「Ｎ」）、「モニター」動作を終了し、ボーター処

理Ｓ１３００を完了して次の再同期検知処理Ｓ１１１を実行する。

＜出力異常処理のフローチャート（図２４）＞

図２４は入出力同期部３１０の実行結果決定処理Ｓ１１６の一部分である、出力異常処理Ｓ１４００の一例を示すフローチャートである。出力異常処理Ｓ１４００は、ボーター処理Ｓ１３００のＳ１３１６で端末２へのアプリケーション２００の処理結果送信に異常が認められた場合に起動される。

[0162] 出力異常処理Ｓ１４００はボーター処理Ｓ１３００のリトライ処理であるため、ほぼボーター処理Ｓ１３００と同様の処理を実行する。しかしリトライ処理であるため、「リーダー」計算機ノード１００、「モニター」計算機ノード１００共に入出力決定処理カウンタ３１５の値は変更しない。

[0163] 入出力同期部３１０は出力異常処理Ｓ１４００を開始すると、初めに決定出力管理テーブル３２１から「リーダー」計算機ノード１００（ボーター処理Ｓ１３００で「リーダー」と判定された計算機ノード１００）の情報を削除する（Ｓ１４０１）。

[0164] 次に入出力同期部３１０は、更新された決定出力管理テーブル３２１の順位３２１Ｂ及び計算機ノード識別子３２１Ｃの値より、自計算機ノード１００が次の「リーダー」であるかを判定する（Ｓ１４０２およびＳ１４０３）。

[0165] 自計算機ノード１００が「リーダー」であると判定される場合（Ｓ１４０３の結果が「Ｙ」）、「リーダー」として動作するために、入出力同期部３１０はＳ１４０５を実行する。自計算機ノード１００が「リーダー」であると判定されない場合（Ｓ１４０３の結果が「Ｎ」）、「モニター」として動作するために、入出力同期部３１０はＳ１４０４を実行する。

[0166] 「リーダー」計算機ノード１００としての動作（Ｓ１４０５からＳ１４１０までの処理、及び再同期検知処理Ｓ１１１の起動）は、ボーター処理Ｓ１３００の「リーダー」動作（Ｓ１３０４からＳ１３０９までの処理、及び再同期検知処理Ｓ１１１の起動）と基本的に同様であるが、Ｓ１４０６とＳ１３０５のみ異なる。出力異常処理Ｓ１４００では入出力決定処理カウンタ３

１５値は変更しない（Ｓ１４０６の点線枠は実行しないことを意味する）。

[0167] 「モニター」計算機ノード１００としての動作（Ｓ１４０４以降の処理）も、ボーター処理Ｓ１３００の「モニター」動作（Ｓ１３０３からＳ１３１７までの処理、及び出力異常処理Ｓ１４００または再同期検知処理Ｓ１１１の起動）と基本的に同様であるが、Ｓ１４１１とＳ１３１２のみ異なる。出力異常処理Ｓ１４００では入出力決定処理カウンタ３１５値は変更しない（Ｓ１４１１の点線枠は実行しないことを意味する）。

＜再同期検知処理および再同期処理のフローチャート（図２５）＞

図２５は再同期検知処理Ｓ１１１および再同期処理Ｓ１１８の一例を示すフローチャートである。再同期検知処理Ｓ１１１は同期システムとして協調動作している系の中に居る計算機ノード１００の入出力同期部３１０が、協調動作の系への参入を要求する計算機ノード１００を検出し、参入を受け入れる処理を行う。再同期処理Ｓ１１８は同期システムとして協調動作している系の外に居る計算機ノード１００の入出力同期部３１０が、協調動作の系への参入を要求し、参入を果たすまでの処理を行う。協調動作の系への参入を要求することを「再同期要求」と称する。

[0168] 再同期検知処理Ｓ１１１と再同期処理Ｓ１１８はそれぞれがREJOIN-REQのノード間通信パケット（詳細は図２６に示す）、REJOIN-ACKのノード間通信パケット（詳細は図２６に示す）、JOINEDのノード間通信パケット（詳細は図２６に示す）を送受信することで、歩調を合わせて処理を進める。

[0169] 図２６に一例として示す「REJOIN-REQのノード間通信パケット」は、再同期要求を行う計算機ノード１００が送信するノード間通信パケット６００の例である。計算機ノード処理状態６０３である計算機ノード処理状態テーブル７００の処理状態識別子７０１と処理状態７０４には「REJOIN-REQ」が設定される。入出力決定処理カウンタ６０２は初期値（例えば「０」）のままで、入出力決定処理カウンタ３１５値は設定されない。

[0170] 図２６に一例として示す「REJOIN-ACKのノード間通信パケット」は再同期要求を受領した計算機ノード１００が要求側に返信するノード間通信パケッ

ト６００の例である。計算機ノード処理状態６０３である計算機ノード処理状態テーブル７００の処理状態識別子７０１と処理状態７０４には「REJOIN-ACK」が設定される。また再同期要求を受領した計算機ノード１００が保持する入出力決定処理カウンタ３１５値が入出力決定処理カウンタ６０２に設定され、決定入力管理テーブル８００情報と外部入力パケット記憶部３０１の外部入力パケット管理テーブル５５０情報が計算機ノード処理状態個別詳細情報６０４に設定される。

[0171] 図２６に一例として示す「JOINEDのノード間通信パケット」は「REJOIN-ACKのノード間通信パケット」を受取った計算機ノード１００が同期システムとして協調動作している系への参入を果たした事を、他の計算機ノード１００に通知するために送信するノード間通信パケット６００の例である。計算機ノード処理状態６０３である計算機ノード処理状態テーブル７００の処理状態識別子７０１と処理状態７０４には「JOINED」が設定される。

[0172] 図２５に戻って、再同期検知処理Ｓ１１１の詳細を示す。

[0173] 入出力同期部３１０は再同期検知処理Ｓ１１１を開始すると、初めに決定出力管理テーブル３２１の出力同期時計算機ノード総数３２１Ａで、計算機ノード構成管理テーブル３１７の入力同期時計算機ノード総数３１７Ａを更新する（Ｓ１１１１）。

[0174] 次に入出力同期部３１０は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の処理状態識別子７０１を「REJOIN-WAIT」に設定し、処理状態７０４を「REJOIN-WAIT」に設定することで、再同期要求を受付ける処理状態になる（Ｓ１１１２）。

[0175] 次に入出力同期部３１０は、他の計算機ノード１００からの再同期要求である、再同期処理Ｓ１１８のＳ１１８３により送信された「REJOIN-REQのノード間通信パケット」の到着を期待して、タイムアウトまでＷａｉｔする（Ｓ１１１３）。

[0176] 再同期要求が検出されなかった場合（Ｓ１１１４の結果が「Ｎ」）、再同期検知処理Ｓ１１１を終了して入力同期処理を開始するために入出力同期部

３１０はＳ１１２を起動する。再同期要求を検出した場合（Ｓ１１１４の結果が「Ｙ」）、再同期要求に応答するために入出力同期部３１０はＳ１１１５を起動する。

[0177] Ｓ１１１５では、入出力同期部３１０は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の処理状態識別子７０１を「REJOIN-ACK」に設定し、処理状態７０４を「REJOIN-ACK」に設定する。計算機ノード１００は再同期要求に応答する計算機ノード処理状態に遷移する。

[0178] 次に入出力同期部３１０は、「REJOIN-ACKのノード間通信パケット」を作成するために、入出力決定処理カウンタ３１５値をノード間通信パケット６００の入出力決定処理カウンタ６０２に設定し、計算機ノード処理状態テーブル７００情報を計算機ノード処理状態６０３に設定する（Ｓ１１１６）。また決定入力管理テーブル８００情報と外部入力パケット記憶部３０１の外部入力パケット管理テーブル５５０情報を計算機ノード処理状態個別詳細情報６０４に設定する（Ｓ１１１７）。

[0179] 作成した「REJOIN-ACKのノード間通信パケット」はノード間通信Ｉ／Ｆ１０５経由でマルチキャストで送信される。当該「REJOIN-ACKのノード間通信パケット」は、再同期要求を行った計算機ノード１００で処理中の、再同期処理Ｓ１１８のＳ１１８５で受信されることを期待する。（Ｓ１１１８）

次に入出力同期部３１０は、再同期要求を行った計算機ノード１００の再同期処理Ｓ１１８のＳ１１８１０から「JOINEDのノード間通信パケット」が到着することを期待して、タイムアウトまでＷａｉｔする（Ｓ１１１９）。

[0180] 「JOINEDのノード間通信パケット」が受信されなかった場合（Ｓ１１２０の結果が「Ｎ」）、再同期要求を行った計算機ノード１００の再同期処理Ｓ１１８は失敗したと判断して、入出力同期部３１０は再同期検知処理Ｓ１１１を終了して入力同期処理を開始するためにＳ１１２を起動する。

[0181] 「JOINEDのノード間通信パケット」が受信された場合（Ｓ１１２０の結果が「Ｙ」）、入出力同期部３１０はＳ１１２１で計算機ノード構成管理テーブル３１７の入力同期時計算機ノード総数３１７Ａを更新した後、再同期検

知処理Ｓ１１１を終了して入力同期処理を開始するためにＳ１１２を起動する。

[0182] 再同期処理Ｓ１１８の詳細を示す。

[0183] 再同期要求を行う計算機ノードＡ１００の入出力同期部Ａ３１０は、初めに処理状態記憶部Ａ３１１の計算機ノード処理状態テーブルＡ７００の処理状態識別子Ａ７０１に「REJOIN-REQ」を設定し、処理状態Ａ７０４に「REJOIN-REQ」を設定設定することで、再同期要求を行う処理状態になる（Ｓ１１８１）。

[0184] 次に入出力同期部Ａ３１０は、「REJOIN-REQのノード間通信パケット」を作成するために、ノード間通信パケットＡ６００の計算機ノード処理状態Ａ６０３に計算機ノード処理状態テーブルＡ７００情報を設定する（Ｓ１１８２）。

[0185] 作成した「REJOIN-REQのノード間通信パケット」はノード間通信Ｉ／Ｆ１０５経由でマルチキャストで送信される。当該「REJOIN-REQのノード間通信パケット」は協調動作している系の中に居る計算機ノードＡ１００が実行している再同期検知処理Ｓ１１１のＳ１１１４で受信されることを期待する。（Ｓ１１８３）

次に再同期要求に対する「REJOIN-ACKのノード間通信パケット」が返信されてくることを期待して、入出力同期部Ａ３１０はタイムアウトまでＷａｉｔする（Ｓ１１８４）。

[0186] 「REJOIN-ACKのノード間通信パケット」が受信されない場合（Ｓ１１８５の結果が「Ｎ」）、再同期要求をリトライするため、入出力同期部Ａ３１０はＳ１１８１を起動する。「REJOIN-ACKのノード間通信パケット」が受信される場合（Ｓ１１８５の結果が「Ｙ」）、入出力同期部Ａ３１０はＳ１１８６を起動する。

[0187] Ｓ１１８６では入出力同期部Ａ３１０は、「REJOIN-ACKのノード間通信パケット」の入出力決定処理カウンタＡ６０２値を入出力決定処理カウンタＡ３１５に設定し、他計算機ノードＡ１００と処理周期を一致させる（Ｓ１１８６）。

[0188] 次に入出力同期部Ａ３１０は、「REJOIN-ACKのノード間通信パケット」の計算機ノード処理状態個別詳細情報Ａ６０４に設定された決定入力管理テーブル

８００情報を取得して自計算機ノード１００の決定入力記憶部３１４の決定入力管理テーブル８００に設定し、同じく外部入力パケット管理テーブル５５０情報を取得して外部入力パケット記憶部３０１の外部入力パケット管理テーブル５５０に設定する（Ｓ１１８７）。

[0189] 次に入出力同期部３１０は、処理状態記憶部３１１の計算機ノード処理状態テーブル７００の処理状態識別子７０１を「JOINED」に更新し、処理状態７０４を「JOINED」に更新して、自計算機ノード１００は再同期が完了した計算機ノード処理状態に遷移する（Ｓ１１８８）。

[0190] 次に入出力同期部３１０は、「JOINEDのノード間通信パケット」を作成するために、ノード間通信パケット６００の計算機ノード処理状態６０３に計算機ノード処理状態テーブル７００情報を設定する（Ｓ１１８９）。

[0191] 作成した「JOINEDのノード間通信パケット」はノード間通信Ｉ／Ｆ１０５経由でマルチキャストで送信される。当該「JOINEDのノード間通信パケット」は「REJOIN-ACKのノード間通信パケット」を返信した計算機ノード１００が実行している再同期検知処理Ｓ１１１のＳ１１２０で受信されることを期待する。（Ｓ１１９０）

「JOINEDのノード間通信パケット」を送信した後、入出力同期部３１０は再同期処理Ｓ１１８を完了して入力同期処理を開始するためにＳ１１２を起動する。

＜ケースを想定した実施態様＞

以降の説明では、処理対象入力パケット選定処理Ｓ１１４（入力多数決処理Ｓ６００までの一連の処理）と実行結果決定処理Ｓ１１６（および出力異常処理Ｓ１４００までの一連の処理）について、想定されるケースを当てはめることで、これまでの説明を補足する。具体的には、計算機ノード処理状態の経過に伴う計算機ノード１００の内部に保持される各テーブル状態や計算機ノード１００間で送受信されるノード間通信パケット６００情報内容の変化を示す。

[0192] 補足説明における同期システム１は、三台の計算機ノード１００Ａ、計算

機ノード１００Ｂ、計算機ノード１００Ｃから構成されるものとするがこれは一例に過ぎない。

[0193] 処理対象入力パケット選定処理Ｓ１１４（入力多数決処理Ｓ６００までの一連の処理）については、二つのケースを当てはめて補足説明する。実行結果決定処理Ｓ１１６（出力異常処理Ｓ１４００までの一連の処理）については、一つケースを当てはめて補足説明する。

[0194] 処理対象入力パケット選定処理Ｓ１１４の一つ目の補足説明では、計算機ノード１００Ｂにだけ端末２からの外部入力パケット５００の到達が遅延（「外部遅延」と称する）したケース（以下ケースＡと称す）を想定し、この結果一回目の外部入力パケット５００の情報の取り出しでは「多数決同意入力パケット」が決定できず、二回目の外部入力パケット５００の情報の取り出しで「多数決同意入力パケット」が決定できて、以降の処理が実行される例を説明する。

[0195] 処理対象入力パケット選定処理Ｓ１１４の二つ目の補足説明では、計算機ノード１００Ｂへの外部遅延の発生に加えて、計算機ノード１００Ｃの処理開始契機が他の計算機ノード１００Ａ、１００Ｂより遅れる（「内部遅延」と称する）ケース（以下ケースＢと称す）を想定し、この結果一回目の外部入力パケット５００の情報の取り出しにおける「多数決同意入力パケット」の決定処理には計算機ノード１００Ｃは参加せず、なおかつ計算機ノード１００Ａ、１００Ｂの間でも「多数決同意入力パケット」の決定が出来ず、二回目の外部入力パケット５００の情報の取り出しで「多数決同意入力パケット」が決定出来て、以降の処理が実行される例を説明する。

[0196] 実行結果決定処理Ｓ１１６の補足説では、計算機ノード１００Ｃの処理開始契機が遅れるケース（以下ケースＣと称す）を想定し、この結果アプリケーション２００の実行結果を計算機ノード１００Ａ、１００Ｂ、１００Ｃの間で共有させる処理（計算機ノード処理状態が「OUTPUT」）に計算機ノード１００Ｃは途中参加することとなり、Ｓ１１６～Ｓ１４００の処理が実行される例を説明する。

＜外部遅延発生時の処理対象入力パケット選定処理（ケースＡ）＞

ケースＡにおける補足説明は図２７、図２８、図２９、図３０、図３１、図３２、図３３を用いて行う。

[0197] 図２７は各計算機ノード１００A、１００B、１００Cの処理シーケンスの一例を示す。時間Ｔ１は処理対象入力パケット選定処理のＳ３００からＳ６０２に対応し、時間Ｔ２は次の処理周期におけるＳ３００からＳ５０９に対応し、時間Ｔ３は時間T２の処理結果を受けて実行されるＳ６０１からＳ６１１に対応する。また送信９０１はＳ４０２に対応し、ＴＯ９０２はＳ４０３に対応し、リスト作成９０３はＳ４０４からＳ５０８に対応し、ＴＯ９０４はＳ５０９に対応し、送信９０５はＳ４０２に対応し、ＴＯ９０６はＳ４０３に対応し、リスト作成９０７はＳ４０４からＳ５０８に対応し、ＴＯ９０８はＳ５０９に対応し、多数決定９０９はＳ６０１に対応し、状態更新９１０はＳ６０４からＳ６０９に対応し、送信９１１はＳ６１１に対応する。

[0198] 図２８は各計算機ノードの外部入力パケット管理テーブル、計算機ノード処理状態テーブル、入出力決定処理カウンタの時間別の内容を示す。列９１２は計算機ノード１００Ａの状態であり、構成列のＳＥQ９１５は外部入力パケット管理テーブル５５０のシーケンス番号５０２を示し、ＷＳ９１６は計算機ノード処理状態テーブル７００の入力同期ステップ７１０を示し、ＳＴ９１７は計算機ノード処理状態テーブル７００の処理状態７０４を示し、ＣＮＴ９１８は入出力決定処理カウンタ３１５を示す。列９１３は計算機ノード１００Ｂの状態を示し、列９１４は計算機ノード１００Ｃの状態を示し、構成列は計算機ノードAの情報と同様である。行９２０は時間Ｔ１における状態を示し、行９２１は時間Ｔ２における状態を示し、行９２２は時間Ｔ３における状態を示す。

[0199] 図２９は図２７に示す「送信」に対する各計算機ノードのノード間通信パケット６００の時間別内容を示す。行９４０は時間Ｔ１の送信９０１に対するノード間通信パケット６００内容を示し、構成行の行９４３は計算機ノード１００Ａの送信内容を示し、行９４４は計算機ノード１００Ｂの送信内容

を示し、行９４５は計算機ノード１００Ｃの送信内容を示す。行９４１は時間Ｔ２の送信９０５に対するノード間通信パケット６００内容を示し、行９４２は時間Ｔ３の送信９１１に対するノード間通信パケット６００内容を示し、構成行は行９４０と同様である。

[0200] 図３０は図２７に示す「リスト作成」に対する各計算機ノードのノード間通信パケット６００の時間別内容を示す。行９６０は時間Ｔ１のリスト作成９０３に対するノード間通信パケット６００内容を示し、これは「送信」処理にて送受信したノード間通信パケット６００を結合した「受信結果リスト」を表す。構成行の行９６２は計算機ノード１００Ａの「送信」内容を示し、行９６３は計算機ノード１００Ｂの「送信」内容を示し、行９６４は計算機ノード１００Ｃの「送信」内容を示す。行９６１は時間Ｔ２のリスト作成９０７に対するノード間通信パケット６００内容を示し、構成行は行９６０と同様である。

[0201] 図３１は任意の計算機ノードの計算機ノード構成管理テーブル３１７の時間別内容を示す。時間別内容は、計算機ノード１００A、１００B、１００Cにおいて同一である。行９８０は時間Ｔ１における計算機ノード構成管理テーブル３１７の計算機ノード識別子６０１列、入出力決定処理カウンタ６０２列、計算機ノード処理状態６０３列の状態を示す。行９８１は時間Ｔ２における計算機ノード構成管理テーブル３１７の保持情報を示し、行９８２はＴ３における計算機ノード構成管理テーブル３１７の保持情報を示す。

[0202] 図３２は各計算機ノードの決定入力管理テーブル８００の内容を示す。図３３は計算機ノード１００Ｂの決定済み入力管理テーブル８１０の内容を示す。

[0203] 次に図２７の処理シーケンスを追って、各計算機ノード１００A、１００B、１００Cの処理状態変化を説明する。

[0204] 時間Ｔ１において、各計算機ノード１００A、１００B、１００Cは図２８の行９２０の状態で処理を開始する。計算機ノード１００Aは、ＷＳ９２４として「２」、ＳＴ９２５として「BEGIN」、ＣＮＴ９２６として「Ｃ０１０」の

値を保持する。計算機ノード１００B、１００Cにおいても同値である。しかし外部入力パケット管理テーブル５５０のに保存している外部入力パケットの情報が異なる。具体的には計算機ノード１００Aはシーケンス番号が「２Ａ－１」と「２Ｂ－１」である外部入力パケット５００を受信して外部入力パケット管理テーブル５５０に当該パケットの情報を登録しているのに対し（枠９３０）、計算機ノード１００Bには外部入力パケット５００が到達していないので外部入力パケット管理テーブル５５０には外部入力パケットの情報が登録されておらず（枠９３１）、計算機ノード１００Ｃはシーケンス番号が「２Ｃ－１」である外部入力パケットのみを受信しているので、当該パケットの情報のみを外部入力パケット管理テーブル５５０に保持している（枠９３２）。尚、外部入力パケット５００の情報が登録されていない状態は「０」で示す。

[0205] 送信９０１では、各計算機ノード１００Ａ、１００Ｂ、１００Ｃは各々、自身が外部ネットワーク３から取得（受信）した外部入力パケット５００の情報をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定して送出する。送出されたノード間通信パケット６００の内容は図２９の時間Ｔ１に示す（行９４０）。計算機ノード１００ＡはＷＳ９２４の値に従い二つの外部入力パケット５００（シーケンス番号「２Ａ－１」と「２Ｂ－１」の外部入力パケット）の情報を取得してノード間通信パケット６００に設定する（行９４３）。計算機ノード１００Ｂ、１００Ｃも二つの外部入力パケット５００の情報を取得しようとする。しかし、計算機ノード１００Ｂは外部入力パケットを受信していないため外部入力パケット５００の情報無しを示す「０」を二つノード間通信パケット６００に設定し（行９４４）、計算機ノード１００Ｃは取得できた一つの外部入力パケット５００（シーケンス番号「２Ｃ－１」の外部入力パケット）の情報ともう一つは「０」をノード間通信パケット６００に設定する（行９４５）。

リスト作成９０３では送信９０１で送受信したノード間通信パケット６００を結合して「受信結果リスト」を作成し、各計算機ノード１００Ａ、１０

０Ｂ、１００Ｃ間で送受信することで計算機ノード間で保持する「受信結果リスト」内容を一致化させる。送受信されるノード間通信パケット６００の内容は図３０の時間Ｔ１に示す（行９６０）。

[0206] 時間Ｔ１内で実行されるＳ６０１では「多数決同意入力パケット」の有無を判定するが、枠９６５には「多数決同意入力パケット」は存在しないため、各計算機ノード１００Ａ、１００Ｂ、１００ＣはＳ６０２で計算機ノード処理状態テーブル７００の入力同期ステップ７１０WSの値を「３」に変更して、次の周期（時間Ｔ２）のＳ３００を起動する。この時の計算機ノード構成管理テーブル３１７の状態を図３１の時間Ｔ１に示す（行９８０）。

[0207] 時間Ｔ２開始状態の各計算機ノード１００A、１００B、１００Cの状態を図２８の時間Ｔ２に示す（行９２１）。計算機ノード１００AのＷＳ９２４は、時間Ｔ１におけるＳ６０２の実行により、値が「３」に変化しているが、これは計算機ノード１００B、１００Cも同様である。時間Ｔ１の処理を実行中に新たな外部入力パケット５００が到着し、時間Ｔ２開始時点で計算機ノード１００Aにはシーケンス番号「２Ｃ－１」の外部入力パケット５００の情報が外部入力パケット管理テーブル５５０に新たに追加され（枠９３３）、計算機ノード１００Bの外部入力パケット管理テーブル５５０にはシーケンス番号「２Ｂ－１」の外部入力パケット５００の情報が追加され（枠９３４）、計算機ノード１００Cの外部入力パケット管理テーブル５５０にはシーケンス番号「２Ｂ－１」の外部入力パケット５００の情報が追加される（枠９３５）。

[0208] 送信９０５では各計算機ノード１００A、１００B、１００Cはそれぞれ三つの外部入力パケット５００の情報の取得を試みて、取得された外部入力パケット５００の情報をノード間通信パケット６００に設定して送出する。送出されるノード間通信パケット６００の内容を　図２９の時間Ｔ２に示す（行９４１）。入力同期ステップ７１０の値の変化が計算機ノード処理状態６０３の付加情報７０２に反映され（枠９３６）、計算機ノード処理状態個別詳細情報６０４には三つ分の外部入力パケット５００の情報が設定される（枠

９３７）。

[0209] またこの時の計算機ノード構成管理テーブル３１７の状態を図３１の時間Ｔ２に示す（行９８１）。付加情報７０２の値が「３」に変更される（枠９８３）。

[0210] リスト作成９０７では送信９０５で送受信されたノード間通信パケット６００の内容に基づき各計算機ノードが「受信結果リスト」を作成し、計算機ノード１００Ａ、１００Ｂ、１００Ｃ間で「受信結果リスト」を送受信することで「受信結果リスト」の内容を一致させる。この時に送受信されるノード間通信パケット６００の内容を図３０の時間Ｔ２に示す（行９６１）。Ｓ６０１での「多数決同意入力パケット」の有無の判定において、時間Ｔ２では、シーケンス番号が「２Ｂ－１」と「２Ｃ－１」の外部入力パケットが「多数決同意入力パケット」として判定される（枠９６６）。「多数決同意入力パケット」を確認したことにより、時間Ｔ３の処理へと移る。

[0211] 時間Ｔ３での多数決定９０９では「多数決同意入力パケット」としてシーケンス番号が「２Ｂ－１」と「２Ｃ－１」の外部入力パケット５００が選出される。

[0212] 状態更新９１０では各計算機ノードは「多数決同意入力パケット」を基に決定入力管理テーブル８００と入出力決定処理カウンタ３１５の更新を行う。各計算機ノード１００A、１００B、１００Cはシーケンス番号「２Ｂ－１」と「２Ｃ－１」の外部入力パケット５００の情報を決定入力管理テーブル８００に設定する。この時の決定入力管理テーブル８００内容は図３２に示す。また同時にシーケンス番号「２Ｂ－１」と「２Ｃ－１」の外部入力パケット５００の情報を外部入力パケット管理テーブル５５０から削除する。「多数決同意入力パケット」の内、シーケンス番号が「２Ｂ－１」の外部入力パケット５００を外部ネットワーク３経由で自ら受信していない計算機ノード１００Bは、決定済み入力管理テーブル８１０にシーケンス番号が「２Ｂ－１」の外部入力パケット５００の情報をを設定する。この時の決定済み入力管理テーブル８１０の内容は図３３に示す。また各計算機ノード１００A、１０

０B、１００Cは入出力決定処理カウンタ３１５の値を「Ｃ０１１」に更新し、これが計算機ノード構成管理テーブル３１７に反映される（図３１の枠９８４）。状態更新９１０により、計算機ノード１００A、１００B、１００Cの状態は、図２８の時間Ｔ３（行９２２）の様に変化する。

[0213] 送信９１１では各計算機ノード１００A、１００B、１００Cがそれぞれ決定したアプリケーション２００の処理対象となる外部入力パケット５００の情報を送受信して、各計算機ノード１００A、１００B、１００C間で共有する。この時送受信されるノード間通信パケット６００の内容は、図２９の時間Ｔ３（行９４２）に示す。

＜処理対象入力パケット選定処理の内部遅延発生の実行例（ケースB）＞

説明は図３４、図３５、図３６、図３７、図３８、図３９、図４０、図４１を用いて行う。

[0214] 図３４は処理対象入力パケット選定処理の内部遅延発生の実行例における各計算機ノード１００A、１００B、１００Cの処理シーケンスを示す。時間Ｔ１は計算機ノード１００A、１００Bの処理対象入力パケット選定処理のＳ３００からＳ６０２の実行時間に対応し、時間Ｔ１’計算機ノード１００Cの処理対象入力パケット選定処理のＳ３００からＳ６０２の実行時間に対応する。時間Ｔ２はその次の実行周期におけるＳ３００からＳ５０９の実行時間に対応し、時間Ｔ３は時間T2の処理結果に基き実行されるＳ６０１からＳ６１１の実行時間に対応する。また送信１００１および送信１０１２はＳ４０２に対応し、ＴＯ１００２およびＴＯ１０１３はＳ４０３に対応し、リスト作成１００３はＳ４０４からＳ５０８に対応し、ＴＯ１００４はＳ５０９に対応し、送信１００５はＳ４０２に対応し、ＴＯ１００６はＳ４０３に対応し、リスト作成１００７はＳ４０４からＳ５０８に対応し、ＴＯ１００８はＳ５０９に対応し、多数決定１００９はＳ６０１に対応し、状態更新１０１０はＳ６０４からＳ６０９に対応し、送信１０１１はＳ６１１に対応する。

[0215] 図３５はケースBにおける各計算機ノードの外部入力パケット管理テーブル、計算機ノード処理状態テーブル、入出力決定処理カウンタの時間別内容を

示す。列および行の構成は、図２８と同様である。

[0216] 図３６はケースBにおいて計算機ノード間で各計算機ノードが受信した外部入力パケット５００の内容を交換するために送受信されるノード間通信パケット６００の時間別内容を示す。列構成と行１０４２と行１０４３の行構成は図２９と同様であるが、時間Ｔ１および時間Ｔ１’ の行構成のみ異なっている。行１０４０は時間Ｔ１の送信１００１に対する計算機ノード１００Ａ、１００Ｂのノード間通信パケット６００内容を示し、行１０４１は時間Ｔ１’ の送信１０１２に対する計算機ノード１００Cのノード間通信パケット６００内容を示す。

[0217] 図３７はケースBにおいて計算機ノード間で各計算機ノードが作成したリストを交換するために送受信されるノード間通信パケット６００の時間別内容を示す。列構成と行１０６１の行構成は図３０と同様であるが、行１０６０の行構成が異なる。行１０６０は時間Ｔ１のリスト作成１００３に対する計算機ノード１００Ａ、１００Ｂのノード間通信パケット６００内容を示す。行１０６４の空行は、計算機ノード１００Ｃがノード間通信を行っていないことを意味する。

[0218] 図３８はケースBにおける計算機ノード１００Ａ、１００Ｂの計算機ノード構成管理テーブル３１７の時間別内容を示す。列構成と行１０８１と行１０８２の行構成は図３１と同様であるが、行１０８０の行構成が異なる。行１０８０は時間Ｔ１における計算機ノード構成管理テーブル３１７の内容を表し、計算機ノード１００Ａ、１００Ｂの情報のみ保持することを示す。行１０８５の空行は、計算機ノード１００Ｃが計算機ノード構成管理テーブル内に情報を保持していないことを意味する。

[0219] 図３９はケースBにおける計算機ノード１００Ｃの計算機ノード構成管理テーブル３１７の時間別内容を示す。列構成と行１０９１と行１０９２の行構成は図３１と同様であるが、行１０９０の行構成が異なる。行１０９０は時間Ｔ１’ における計算機ノード構成管理テーブル３１７の内容を表し、計算機ノード１００Ｃの情報のみ保持することを示す。行１０９３および行１０

９４の空白は、計算機ノード１００Ａ、１００Ｂが計算機ノード構成管理テーブルに情報を保持していないことを意味する。

[0220] 図４０はケースBにおける各計算機ノードの決定入力管理テーブル８００の内容を示す。図４１はケースBにおける計算機ノード１００Ｂの決定済み入力管理テーブル８１０の内容を示す。

[0221] 次に図３４の処理シーケンスを追って、各計算機ノード１００A、１００B、１００Cの処理状態変化を説明する。

[0222] 時間Ｔ１において、各計算機ノード１００A、１００B、１００Cは図２８の行１０２０の状態で処理を開始する。計算機ノード１００AはＷＳ１０２５は「２」、ＳＴ１０２６は「BEGIN」、ＣＮＴ１０２７は「Ｃ０１０」の値を保持する。計算機ノード１００B、１００Cにおいても同値である。しかし外部入力パケット管理テーブル５５０に保存されている情報の内容が異なり、計算機ノード１００Aはシーケンス番号「２Ａ－１」と「２Ｂ－１」の外部入力パケットの情報を保持し（枠１０３１）、計算機ノード１００Ｂはシーケンス番号「２Ｃ－１」の外部入力パケットの情報のみを保持する（枠１０３２）。外部入力パケット５００の情報を保存していない場合は「０」で示す。計算機ノード１００Ｃは外部入力パケット５００情報を保持していない（枠１０３３）。

[0223] 送信１００１では、計算機ノード１００Ａ、１００Ｂは取得した外部入力パケット５００の情報をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定して送出する。尚時間Ｔ１で計算機ノード１００Cには起動契機となる外部入力パケット５００が到着していないため、計算機ノード１００Ｃは送信１００１は行わない。送出されたノード間通信パケット６００の状態は図３６の時間Ｔ１に示す（行１０４０）。計算機ノード１００ＡはＷＳ１０２５の値に従い二つの外部入力パケット５００の情報（シーケンス番号が「２Ａ－１」と「２Ｂ－１」の外部入力パケットの情報）を取得してノード間通信パケット６００に設定する（行１０４４）。計算機ノード１００Ｂも二つの外部入力パケット５００の情報を取得しようとする

が、取得できるのは１つの外部入力パケット５００だけなので、取得できた一つの外部入力パケット５００の情報（シーケンス番号「２Ｃ－１」の外部入力パケットの情報）ともう一つは「０」を設定する（行１０４５）。

[0224] 時間Ｔ１’ になって、外部入力パケット５００を受信した計算機ノード１００Cは、当該外部入力パケット５００のシーケンス番号を用いて送信１０１２を行う。送出されたノード間通信パケット６００の内容を図３６の時間Ｔ１’ に示す（行１０４６）。

[0225] 計算機ノード１００Ｃが送信１０１２を行っている間に、計算機ノード１００Ａ、１００Ｂは「受信結果リスト」を作成し、計算機ノード間で送信する（リスト作成１００３）。このリスト作成１００３で送受信するノード間通信パケット６００の内容を図３７の時間Ｔ１に示す（行１０６０）。「受信結果リスト」に計算機ノード１００Ｃが取得した外部入力パケット５００の情報は含まれていない（行１０６４）。

[0226] ＴＯ１００４で計算機ノード１００Ａ、１００Ｂは双方が送信１０１２で送信した「受信結果リスト」を取得し、Ｓ６０１で「多数決同意入力パケット」の有無を判定するが、枠１０６５に「多数決同意入力パケット」は存在しないため、計算機ノード１００Ａ、１００ＢはＳ６０２で計算機ノード処理状態テーブル７００の入力同期ステップ７１０の値を「３」に変更して時間Ｔ２内のＳ３００を起動する。

[0227] 一方計算機ノード１００Ｃは、計算機ノード１００Ａ、１００Ｂから送信１００１で送信されるノード間通信パケット６００の受信を期待してＴＯ１０１３を実行するが、代わりにリスト作成１００３で送信された「受信結果リスト」を受信する。そして計算機ノード１００Cは、Ｓ５０４およびＳ５０５で計算機ノード処理状態の不一致を検出し、Ｓ５０６で計算機ノード処理状態テーブル７００のPenalty７０３の値を「１」に更新する。 計算機ノード１００Ａ、１００Ｂから計算機ノード１００Cが受信した「受信結果リスト」中に「多数決同意入力パケット」をが存在しないため、計算機ノード１００CはＳ６０２で計算機ノード処理状態テーブル７００の入力同期ステップ７

１０の値を「３」に変更して時間Ｔ２内のＳ３００を起動する。

[0228] 時間Ｔ１における計算機ノードＡ、ＢのノードＣ構成管理テーブル３１７の内容を図３８に示す（行１０８０）。計算機ノードＡ、ＢはノードＣの情報を計算機ノード構成管理テーブルに保持していない（行１０８５の空行）。また時間Ｔ１’ における計算機ノードＣの計算機ノード構成管理テーブル３１７の状態を図３９に示す（行１０９０）。計算機ノードＣは計算機ノードＡ、ＢのＣの情報は計算機ノード構成管理テーブルに保持していない（行１０９３、行１０９４の空行）。

[0229] 時間Ｔ２の開始時点における各計算機ノードA、B、Cの状態を図３５の時間Ｔ２に示す（行１０２９）。計算機ノードAのＷＳ１０２３は、時間Ｔ１におけるＳ６０２の実行により、値が「３」に変化しているが、これは計算機ノードB、Cも同様である。計算機ノードＣには、時間Ｔ１’ の処理中に到着した新たな外部入力パケット５００の情報（シーケンス番号「２Ａ－１」の外部入力パケットの情報）が保持される（枠１０３４）。

[0230] 送信１００５では各計算機ノードA、B、Cはそれぞれ三つの外部入力パケット５００情報の取得を試みて、取得された外部入力パケット５００情報をノード間通信パケット６００に設定して送出する。送出されるノード間通信パケット６００の状態を図３６の時間Ｔ２に示す（行１０４２）。計算機ノードＣが送出したノード間通信パケット６００には、時間Ｔ１’ のＳ５０６で更新されたPenalty７０３の値「１」が保持される（枠１０４７）。

[0231] リスト作成１００７では送信１００５で送受信されたノード間通信パケット６００から「受信結果リスト」を作成し、各計算機ノードＡ、ＢＢ、ＣＣ間で「受信結果リスト」内容を一致化させる。この時に送受信されるノード間通信パケット６００の内容を図３７の時間Ｔ２に示す（行１０６１）。Ｓ６０１での「多数決同意入力パケット」の有無を判定におい

て、時間Ｔ２では、シーケンス番号「２Ａ－１」の外部入力パケットが「多数決同意入力パケット」として判定される（枠１０６５と枠１０６６）。「多数決同意入力パケット」を確認したことにより、時間Ｔ３の処理へと移る。

[0232] 時間Ｔ２における計算機ノード１００Ａ、１００Ｂの計算機ノード構成管理テーブル３１７の状態を図３８に示す（行１０８１）。計算機ノード１００Ａ、１００Ｂは同期した計算機ノード１００Ｃの情報を計算機ノード構成管理テーブルに反映させる（行１０８６）。また計算機ノード１００Ｃの計算機ノード構成管理テーブル３１７の状態を図３９に示す（行１０９１）。計算機ノード１００Ｃは計算機ノード１００Ａ、１００Ｂの情報を新たに計算機ノード構成管理テーブルに保持する（行１０９６と行１０９７）。

[0233] 時間Ｔ３での多数決定９０９では「多数決同意入力パケット」としてシーケンス番号が「２Ａ－１」の外部入力パケットを選出する。状態更新９１０では「多数決同意入力パケット」を基に決定入力管理テーブル８００と入出力決定処理カウンタ３１５の更新を行う。各計算機ノード１００A、１００B、１００Cはシーケンス番号が「２Ａ－１」の外部入力パケット５００の情報を決定入力管理テーブル８００に設定する。この時の決定入力管理テーブル８００内容は図４０に示す。また同時にシーケンス番号が「２Ａ－１」の外部入力パケット５００の情報を外部入力パケット管理テーブル５５０から削除する。

[0234] 「多数決同意入力パケット」となったシーケンス番号「２Ａ－１」の外部入力パケット５００を自ら受信していない計算機ノード１００Bは、決定済み入力管理テーブル８１０にシーケンス番号「２Ａ－１」の外部入力パケット５００の情報をを設定する。この時の決定済み入力管理テーブル８１０内容は図４１に示す。また各計算機ノード１００A、１００B、１００Cは入出力決定処理カウンタ３１５の値を「Ｃ０１１」に更新し、これが計算機ノード構成管理テーブル３１７に反映される（図３８の枠１０８７と図３９の枠１０９８）。状態更新９１０により、計算機ノード１００A、１００B、１００Cの

状態は、図３５の時間Ｔ３（行１０３０）の様に変化する。

[0235] 送信９１１では各計算機ノード１００A、１００B、１００Cがそれぞれ決定したアプリケーション２００の処理対象となる外部入力パケット５００情報を送受信して、各計算機ノード１００A、１００B、１００C間で共有する。この時送受信されるノード間通信パケット６００の内容は、図３６の時間Ｔ３（行１０４３）に示す。

＜実行結果決定処理（ケースC）＞

説明は図４２、図４３、図４４、図４５、図４６、図４７、図４８、図４９を用いて行う。

[0236] 図４２は各計算機ノードの処理シーケンスを示す。

[0237] 時間Ｔ１は計算機ノード１００A、１００Bの実行結果決定処理のＳ８０１からＳ９１３に対応し、時間Ｔ１’ は計算機ノード１００Cの実行結果決定処理のＳ８０１からＳ９１３に対応する。時間Ｔ２はＳ１００１からＳ１２０８に対応する。計算機ノード１００Bでは時間Ｔ３はＳ１３０１からＳ１３０８に対応し、時間Ｔ４はＳ１３０９以降の処理に対応する。計算機ノード１００A、１００Ｃでは時間Ｔ３はＳ１３０１からＳ１３１５に対応し、時間Ｔ４はＳ１３１６およびＳ１３１７以降の処理に対応する。また送信１１０１および送信１１０５はＳ８０４に対応し、ＴＯ１１０２およびＴＯ１１０６はＳ８０５に対応し、リスト作成１１０３はＳ８０６からＳ９１２に対応し、ＴＯ１１０４はＳ９１３に対応する。多数決定１１０７はＳ１００１からＳ１１０５に対応し、状態更新１１０８はＳ１１０１からＳ１１０４に対応し、送信１１０９はＳ１１０５に対応し、ＴＯ１１１０はＳ１１０６に対応し、リスト作成１１１９はＳ１１０７からＳ１２０８に対応する。状態更新１１２０はＳ１３０１からＳ１３０５に対応し、送信１１２１はＳ１３０６とＳ１３０７に対応し、ＴＯ１１２２はＳ１３０８に対応し、出力１１２３はＳ１３０９に対応する。状態更新１１２４はＳ１３０１からＳ１３１２に対応し、送信１１２５はＳ１３１３とＳ１３１４に対応し、ＴＯ１１２６はＳ１３１５に対応し、監視１１２７はＳ１３１６とＳ１３１７以降の処理

に対応する。

[0238] 図４３は時間Ｔ１およびＴ１'における送信（１１０１、１１０５）で送受信されるノード間通信パケットの内容を示す。計算機ノード処理状態個別詳細情報６０４には、各計算機ノード１００Ａ、１００Ｂ、１００Ｃで実行したアプリケーション２００の演算結果を保持する。行１１３１は時間Ｔ１の送信１１０１において送受信されたノード間通信パケット６００の内容を示し、具体的には構成行の行１１３３は計算機ノード１００Ａの送信内容を示し、行１１３４は計算機ノード１００Ｂの送信内容を示すが、行１１３５の空行は時間Ｔ１に計算機ノード１００Ｃは送信を行わなかったことを意味する。行１１３２は時間Ｔ１' の送信１１０５において送受信される通信パケット６００の内容を示し、具体的には構成行の行１１３８は計算機ノード１００Ｃの送信内容を示すが、行１１３６と行１１３７の空行は計算機ノード１００Ａ、１００Ｂは送信していないことを意味する。

[0239] 図４４は時間Ｔ２にて送信１１０９で各計算機ノード１００Ａ、１００Ｂ、１００Ｃ間で送受信されるノード間通信パケットの内容を示す。計算機ノード処理状態個別詳細情報６０４には、各計算機ノード１００Ａ、１００Ｂ、１００Ｃが決定した「リーダー選出結果」を保持する。行１１４０は時間Ｔ２の送信１１０９において送受信されるノード間通信パケット６００の内容を示し、具体的には構成行の行１１４１は計算機ノード１００Ａの送信内容を示し、行１１４２は計算機ノード１００Ｂの送信内容を示し、行１１４３は計算機ノード１００Ｃの送信内容を示す。

[0240] 図４５は時間Ｔ３にて送信１１２１および１１２５で各計算機ノードが送受信するノード間通信パケットの内容を示す。

[0241] 行１１５０は時間Ｔ３の送信１１２１および送信１１２５において送受信されるノード間通信パケット６００内容を示し、具体的には構成行の行１１５１は計算機ノード１００Ａの送信内容を示し、行１１５２は計算機ノード１００Ｂの送信内容を示し、行１１５３は計算機ノード１００Ｃの送信内容を示す。

[0242] 図４６は時間Ｔ１のリスト作成１１０３において、計算機ノード１００Ａ、１００Ｂが送受信するノード間通信パケットの内容を示す。行１１６０は時間Ｔ１のリスト作成１１０３において計算機ノード１００Ａ、１００Ｂの間で送受信されるノード間通信パケット６００内容を示し、これは「送信」処理にて送受信したノード間通信パケット６００を結合した「受信結果リスト」を表す。構成行の行１１６１は計算機ノード１００Ａの「送信」内容を示し、行１１６２は計算機ノード１００Ｂの「送信」内容を示す。行１１６３の空行は、計算機ノード１００Ｃからはノード間通信パケット６００を受取っていないため、「受信結果リスト」に計算機ノード１００Cから受信したノード間通信パケットが含まれていないことを意味する。

[0243] 図４７は時間Ｔ２のリスト作成１１１９において、各計算機ノード間で送受信されるノード間通信パケットの内容を示す。行１１７０はリスト作成１１１９において各計算機ノード１００Ａ、１００Ｂ、１００Ｃの間で送受信されるノード間通信パケット６００内容を示し、これは「送信」処理にて送受信したノード間通信パケット６００を結合した「受信結果リスト」を表す。構成行の行１１７１は計算機ノード１００Ａの「送信」内容を示し、行１１７２は計算機ノード１００Ｂの「送信」内容を示し、行１１７３は計算機ノード１００Ｃの「送信」内容を示す。

[0244] 図４８は任意計算機ノードの計算機ノード構成管理テーブルの時間別内容を示す。時間別内容は、計算機ノード１００A、１００B、１００Cにおいて同一である。行１１８０は時間Ｔ１における計算機ノード構成管理テーブル３１７の計算機ノード識別子６０１列、入出力決定処理カウンタ６０２列、計算機ノード処理状態６０３列の状態を示す。行１１８１は時間Ｔ２における計算機ノード構成管理テーブル３１７の保持情報を示し、行１１８２はＴ３における計算機ノード構成管理テーブル３１７の保持情報を示す。入力同期時計算機ノード総数３１７Ａは入力同期完了時に設定された値（図４８中では「３」）であり、出力同期処理中は変更されない。

[0245] 図４９は実行結果決定処理の実行例における任意計算機ノードの各計算機

ノードの決定出力管理テーブルの内容を示す。

[0246] 次に図４２の処理シーケンスを追って、各計算機ノード１００A、１００B、１００Cの処理状態変化を説明する。

[0247] 送信１１０１で計算機ノード１００A、１００Bは選定入力パケット処理Ｓ１１５によるアプリケーション２００の演算結果「Ｘ」をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定して送出する。送出されたノード間通信パケット６００の内容は図４３の時間Ｔ１に示す（行１１３１）。計算機ノード１００Ｃは実行結果決定処理の開始が遅れているため、時間Ｔ１での送信は行っていない（行１１３５の空行）。

[0248] リスト作成１１０３では計算機ノード１００A、１００Bは送信１１０１で送受信したノード間通信パケット６００を結合して「受信結果リスト」を作成し、各計算機ノード１００A、１００B、１００Cに対して送信する。送出されたノード間通信パケット６００の内容を図４６に示す。「受信結果リスト」には、計算機ノード１００A、１００Bから受信した情報は含まれるが（行１１６１と行１１６２）、計算機ノード１００Ｃから受信した情報は含まれない（行１１６３の空行）。

[0249] 計算機ノード１００A、１００BはTO１１０４で双方の「受信結果リスト」を取得し、多数決定１１０７の処理を開始する。

[0250] 一方計算機ノード１００Ｃは時間Ｔ１’ で実行結果決定処理を開始し、送信１１０５でアプリケーション２００の演算結果「Ｘ」をノード間通信パケット６００の計算機ノード処理状態個別詳細情報６０４に設定して送出する。時間Ｔ１’ における送出されたノード間通信パケット６００の内容を図４３の時間Ｔ１’ に示す（行１１３２）。時間Ｔ１’ で送信を行っているは計算機ノード１００Ｃのみで（行１１３８）、計算機ノード１００A、１００Bは送信していない（行１１３７と行１１３８の空行）。

[0251] 計算機ノード１００CはTO１１０６で計算機ノード１００A、１００Bからそれぞれの計算機ノードのアプリケーション２００による演算の演算結果を受信することを期待するが、代りにリスト作成１１０３で送信された「受信

結果リスト」を受信する。Ｓ９０４およびＳ９０５で計算機ノード１００Cは計算機ノード処理状態の不一致を検出し、Ｓ９０６で計算機ノード１００Cha計算機ノード処理状態テーブル７００のPenalty７０３の値を「１」に更新する。計算機ノード１００Ｃは受信した「受信結果リスト」に自身のアプリケーション２００の演算結果を加えて「受信結果リスト」を作成し、多数決定１１０７の処理を開始する。

[0252] 多数決定１１０７では、計算機ノード１００A、１００Bはリスト作成１１０３で共有した「受信結果リスト」を基にアプリケーション２００の演算結果の多数決判定を行う。 計算機ノード１００ＣはＴＯ１１０６で取得、作成した「受信結果リスト」を基に多数決判定を行う。本例では計算機ノード１００A、１００B、１００Ｃのアプリケーション２００の演算結果は全て同一の「Ｘ」であるため、「受信結果リスト」に含まれる全ての計算機ノードが「多数派」として判断される。

[0253] 状態更新１１０８では「受信結果リスト」に含まれる計算機ノードから、リーダー候補となる計算機ノードを選出し、リーダー選出の順位付けを行う。計算機ノード１００A、１００Bが保持する「受信結果リスト」には「計算機ノード１００A、１００B」が含まれているため、何れもＳ１１０２で「計算機ノード１００A、１００B」を選出し、「計算機ノード１００B」を一位、「計算機ノード１００A」を二位と順位付けする（本例では計算機ノード識別子の値の降順で順位付けを行う）。一方計算機ノード１００Ｃが保持する「受信結果リスト」には「計算機ノード１００A、１００B、１００Ｃ」が含まれているため、「計算機ノード１００A、１００B、１００Ｃ」を選出し、「計算機ノード１００B」を一位、「計算機ノード１００A」を二位、「計算機ノード１００Ｃ」を三位と順位付けする。「計算機ノード１００Ｃ」は最も大きな計算機ノード識別子の値を持つが、Penalty７０３に「１」が設定されているため最下位の三位に順位付けされる。

[0254] 送信１１０９では各計算機ノード１００A、１００B、１００Cは状態交信１１０８のリーダー選出結果をノード間通信パケット６００の計算機ノード処

理状態個別詳細情報６０４に設定して送出する。送信されるノード間通信パケット６００の内容を図４４に示す。計算機ノード１００A、１００Bは「計算機ノード１００C」のリーダー選出順位は決定せずに送信する（枠１１４４と枠１１４５）。一方計算機ノード１００Cは自身のリーダー選出順位を付加して送信する（枠１１４６）。

[0255] リスト作成１１１９では計算機ノード１００A、１００B、１００C間で交換したリーダ選出順位を照合して結合した「受信結果リスト」の作成と共有を行う。「受信結果リスト」が設定された、リスト作成１１１９で送受信されるノード間通信パケット６００の内容を図４７に示す。計算機ノード１００A、１００Bが作成した「受信結果リスト」には、計算機ノード１００Cのリーダー選出順位が補完され、反映される（枠１１７４と枠１１７５）。

[0256] 時間Ｔ３以降は、計算機ノード１００A、１００B、１００Cはリーダー選出順位に基き「リーダー」と「モニター」に分かれ、それぞれ異なる処理を行う。計算機ノード１００Bは「リーダー」になり、計算機ノード１００A、１００Cは「モニター」となる。

[0257] 状態更新１１２０で計算機ノード１００Bは、リーダー選出順位により自身を「リーダー」と判断し、計算機ノード処理状態テーブル７００の処理状態７０４に「LEADER」を設定し、入出力決定処理カウンタ３１５の値を変更する。計算機ノード１００A、１００Cについては「モニター」と判断する。

[0258] 一方状態更新１１２４で計算機ノード１００A、１００Cは、リーダー選出順位により自身を「モニター」と判断し、計算機ノード処理状態テーブル７００の処理状態７０４に「MONITOR」を設定し、入出力決定処理カウンタ３１５の値を変更する。計算機ノード１００Bについては「リーダー」と判断する。

[0259] 状態更新１１２０および状態更新１１２４は、自他計算機ノードの情報および「リーダー／モニター」判断結果を計算機ノード構成管理テーブル３１７と決定出力管理テーブル３２１に反映する。時間Ｔ３における計算機ノード構成管理テーブル３１７の内容を図４８の時間Ｔ３（行１１８２）に、決

定出力管理テーブル３２１の内容を図４９に示す。

[0260] 送信１１２１および送信１１２５で、各計算機ノード１００A、１００B、１００Cは自身の計算機ノード処理状態を送信する。送信１１２１および送信１１２５で送信されるノード間通信パケット６００の内容を図４５に示す。各計算機ノード１００A、１００B、１００Cの更新された入出力決定処理カウンタ６０２値（枠１１５４）と、「リーダー／モニター」を示す処理状態７０４（枠１１５５）が通知される。

[0261] 出力１１２３で「リーダー」である計算機ノード１００BはＳ１３０９を実行し、アプリケーション２００の演算結果である信号を外部に送信する。

[0262] 監視１１２７において「モニター」である計算機ノード１００A、１００Cは、Ｓ１３１６およびＳ１３１７を実行して出力１１２３の信号送出を監視することで「リーダー」である計算機ノード１００Bの外部出力の正常／異常を判断し、再同期検知処理Ｓ１１１または出力異常処理Ｓ１４００を実行する。

実施例 2

[0263] 次に別の実施形態として、既存の稼動中の多数決同期システムの系に新たな計算機ノードを加えた場合の動作説明を行う。

[0264] 図５０は本発明の実施の形態の、新規計算機ノード追加における多数決同期システムの全体構成例である。同期システム１Ａは、計算機ノード１００Ａ、１００Ｂ、１００Ｃから構成される、既存稼動中の多数決同期システムの系を示す。計算機ノード１００Ｄは同期システム１Ａに新たに追加される計算機ノードであり、図１の計算機ノード１００Ａと同様のハードウェア構成、ソフトウェア構成を保持する。外部ネットワーク３は計算機ノード１００Ａ、１００Ｂ、１００Ｃおよび計算機ノード１００Ｄと端末２Ａ、２Ｂ、２Ｃを接続し、内部ネットワーク４は計算機ノード１００Ａ、１００Ｂ、１００Ｃおよび計算機ノード１００Ｄの間を接続する。同期システム１Ｂは計算機ノード１００Ｄ追加後の、計算機ノード１００Ａ、１００Ｂ、１００Ｃ、１００Ｄから構成される多数決同期システムの系を示す。

[0265] 次に同期システム１Ａに計算機ノード１００Ｄを追加する時の動作説明を行う。なお動作説明は、外部ネットワーク３と内部ネットワーク４への計算機ノード１００Ｄ接続操作は省略し、接続が完了した状態から始める。

[0266] 計算機ノード１００Ｄで基盤ソフトウェア３００の入出力同期部３１０が動作を開始すると、計算機ノード１００Dの入出力同期部３１０は再同期処理Ｓ１１８を実行し、Ｓ１１８３で「REJOIN-REQのノード間通信パケット」を内部ネットワーク４に接続された計算機ノード１００Ａ、１００Ｂ、１００Ｃに対してマルチキャストで送信する。そしてＳ１１８４で「REJOIN-ACKのノード間通信パケット」の到着を期待してタイムアウトまでＷａｉｔする。

[0267] 計算機ノード１００Ａ、１００Ｂ、１００Ｃは、再同期検知処理Ｓ１１１のＳ１１１４で計算機ノード１００Ｄからの「REJOIN-REQのノード間通信パケット」を検出し、Ｓ１１１８で「REJOIN-ACKのノード間通信パケット」を計算機ノード１００Ｄを含む他の計算機ノードに対してマルチキャストで送信する。そして算機ノード１００Ｄからの「JOINEDのノード間通信パケット」の到着を期待してタイムアウトまでＷａｉｔする。

[0268] 計算機ノード１００Ｄは計算機ノード１００Ａ、１００Ｂ、１００Ｃからの「REJOIN-ACKのノード間通信パケット」を受信すると、Ｓ１１８１０で「REJOIN-REQのノード間通信パケット」を計算機ノード１００Ａ、１００Ｂ、１００Ｃに対してマルチキャストで送信し、入力同期処理を開始するためにＳ１１２を起動する。

計算機ノード１００Ａ、１００Ｂ、１００Ｃは計算機ノード１００Ｄからの「REJOIN-REQのノード間通信パケット」の受信を確認すると、入力同期処理を開始するためにＳ１１２を起動する。

[0269] これ以降、多数決同期システムは計算機ノード１００－Ｄを加えた系（同期システム１－Ｂ）で端末２－Ａ、２－Ｂ、２－Ｃからの外部入力パケット５００を処理する。

符号の説明

[0270] １：同期システム

２A、２B、２C：端末

１００A、１００B、１００C：計算機ノード

１０１：メモリ

１０２：プロセッサ

１０３：記憶装置

１０４：外部通信I/F

１０５：ノード間通信I/F

２００：アプリケーション

３００：基盤ソフトウェア

３０１：外部入力パケット記憶部

３１０：入出力同期部

４００：オペレーティングシステム

５００：外部入力パケット

６００：ノード間通信パケット

請求の範囲

[請求項1] 複数の計算機ノードを有する計算機システムであって、

各計算機ノードは、処理要求を受信する外部通信部と、当該処理要求に従って処理を実行するアプリケーション部と、各計算機ノードのアプリケーション部で実行される処理の他の計算機ノードとの間での同期を制御する同期部と、他の計算機ノードとの間で情報の送受信を実行するノード間通信部とを有し、

前記各計算機ノードの前記同期部は、前記外部通信部が受信した処理要求識別情報が付与された処理要求に、当該計算機ノードの識別情報を送信元計算機ノード識別情報として付与して前記ノード間通信部を介して他の計算機ノードに送信し、

前記各計算機ノードの前記同期部は、前記ノード間通信部を介して他の計算機ノードから当該他の計算機ノードの識別情報が付与された処理要求を受信し、

前記各計算機ノードの前記同期部は、同じ処理要求識別情報が付与された処理要求を外部通信部を経由して受信した計算機ノードの数に基づく第一の所定の判断基準に従って、前記ノード間通信部が他の計算機ノードから受信した処理要求と前記外部通信部が受信した処理要求の中から、前記アプリケーション部に実行させる処理要求を選択し、

前記アプリケーション部は選択された前記処理要求を実行することを特徴とする計算機システム。

[請求項2] 請求項１記載の計算機システムであって、

前記第一の所定の判断基準とは、前記複数の計算機ノードの過半数が同じ処理要求識別情報が付与された処理要求を外部通信部を経由して受信しているか否かであることを特徴とする計算機システム。

[請求項3] 請求項１記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記アプリケーション部が選

択された前記処理要求を実行した実行結果を、前記ノード間通信部を介して他の計算機ノードに送信し、

前記各計算機ノードの前記同期部は、前記ノード間通信部を介して他の計算機ノードから実行結果を受信し、

前記各計算機ノードの前記同期部は、同じ実行結果を取得した計算機ノードの数に基く第二の所定の判断基準に従って、前記ノード間通信部が他の計算機ノードから受信した実行結果と前記アプリケーション部の実行結果の中から、出力結果を選択し、

前記各計算機ノードの前記同期部は、選択された前記出力結果を前記アプリケーション部から実行結果として受け取っていた場合に、選択された前記出力結果をアプリケーション部の実行結果として送信した他の計算機および自計算機ノードの中から、予め定められた基準に従ってリーダ計算機ノード候補を選出して、当該リーダ計算機ノード候補の識別情報を、前記ノード間通信部を介して他の計算機ノードに送信し、

前記各計算機ノードの前記同期部は、前記ノード間通信部を介して、他の計算機ノードから当該他の計算機ノードが選出したリーダ計算機ノード候補の識別情報を受信し、

同じ識別情報で識別される計算機ノードをリーダ計算機ノード候補として選出した計算機ノードの数に応じて、自計算機ノードが選出した前記リーダ計算機ノード候補をリーダ計算機ノードとするか否かを判断し、

自計算機ノードが選出した前記リーダ計算機ノード候補をリーダ計算機ノードとした場合であって、自計算機ノードが前記リーダ計算機ノードに選出された場合には、選択された前記出力結果を出力することを特徴とする計算機システム。

[請求項4] 請求項３記載の計算機システムであって、

前記第二の所定の判断基準とは、前記複数の計算機ノードの過半数

が、同じ実行結果をアプリケーション処理部から受け取っているか否かであることを特徴とする計算機システム。

[請求項5] 請求項３記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードが選出した前記リーダ計算機ノード候補をリーダ計算機ノードとした場合であって、前記自計算機ノードが前記リーダ計算機ノードに選出されなかった場合に、前記リーダ計算機ノードに選出された他の計算機ノードが出力する出力結果を監視し、

前記リーダ計算機ノードに選出された他の計算機が出力した出力結果が、前記自計算機ノードが選択した出力結果とは異なる場合に、前記自計算機ノードが選出した出力結果と同じ出力結果を選択した他の計算機ノードおよび自計算機ノードの中から別のリーダ計算機ノードを選出することを特徴とする計算機システム。

[請求項6] 請求項３記載の計算機システムであって、

前記各計算機ノードの前記同期部は、当該同期部が実行する処理の進捗状況を示す処理状態情報を有しており、

前記各計算機ノードの前記同期部は、前記外部通信部が受信した処理要求に更に前記処理状態情報を付与して前記ノード間通信部を介して前記他の計算機ノードに送信し、

前記各計算機ノードの前記同期部は、前記ノード間通信部を介して他の計算機から当該他の計算機の処理状態情報が付与された処理要求を受信し、

前記各計算機ノードの前記同期部は、前記自計算機ノードの処理状態情報と前記他の計算機から受信した処理状態情報とを比較し、前記自計算機ノードの同期部による処理の進捗の遅れが所定の基準以上である場合に、前記アプリケーション部に実行させる処理要求を選択するための処理を実行せずに、前記他の計算機ノードとの間で同期をとるよう再同期要求を当該他の計算機ノードに送信することを特徴とす

る計算機システム。

[請求項7] 請求項６記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードと同期をとっている計算機ノードの数を管理しており、

前記各計算機ノードの前記同期部は、他の計算機ノードから再同期要求を受信した場合に、前記処理状態情報を当該他の計算機ノードに送信し、当該他の計算機ノードからの応答を受信すると、前記計算機ノードの数を更新することを特徴とする計算機システム。

[請求項8] 請求項６記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードの処理状態情報と前記他の計算機から受信した処理状態情報とを比較し、前記自計算機ノードの同期部による処理の遅れが所定の基準以内である場合には、前記アプリケーション部に実行させる処理要求を選択するための処理を実行することを特徴とするシステム。

[請求項9] 請求項項３記載の計算機システムであって、

前記各計算機ノードの前記同期部は、当該同期部が実行する処理の進捗状況を示す処理状態情報を有しており、

前記各計算機ノードの前記同期部は、前記アプリケーション部の実行結果に更に前記処理状態情報を付与して前記ノード間通信部を介して前記他の計算機ノードに送信し、

前記各計算機ノードの前記同期部は、前記ノード間通信部を介して他の計算機から当該他の計算機の処理状態情報が付与された実行結果を受信し、

前記各計算機ノードの前記同期部は、前記自計算機ノードの処理状態情報と前記他の計算機から受信した処理状態情報とを比較し、前記自計算機ノードの同期部による処理の進捗の遅れが所定の基準以上である場合に、前記出力結果を選択するための処理を実行することなく、前記他の計算機ノードとの間で同期をとるよう再同期要求を当該他

の計算機ノードに送信することを特徴とする計算機システム。

[請求項10] 請求項９記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードと同期をとっている計算機ノードの数を管理しており、

前記各計算機ノードの前記同期部は、他の計算機ノードから再同期要求を受信した場合に、前記処理状態情報を当該他の計算機ノードに送信し、当該他の計算機ノードからの応答を受信すると、前記計算機ノードの数を更新することを特徴とする計算機システム。

[請求項11] 請求項９記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードの処理状態情報と前記他の計算機から受信した処理状態情報とを比較し、前記自計算機ノードの同期部による処理の遅れが所定の基準以内である場合には、前記出力結果を選択するための処理を実行することを特徴とするシステム。

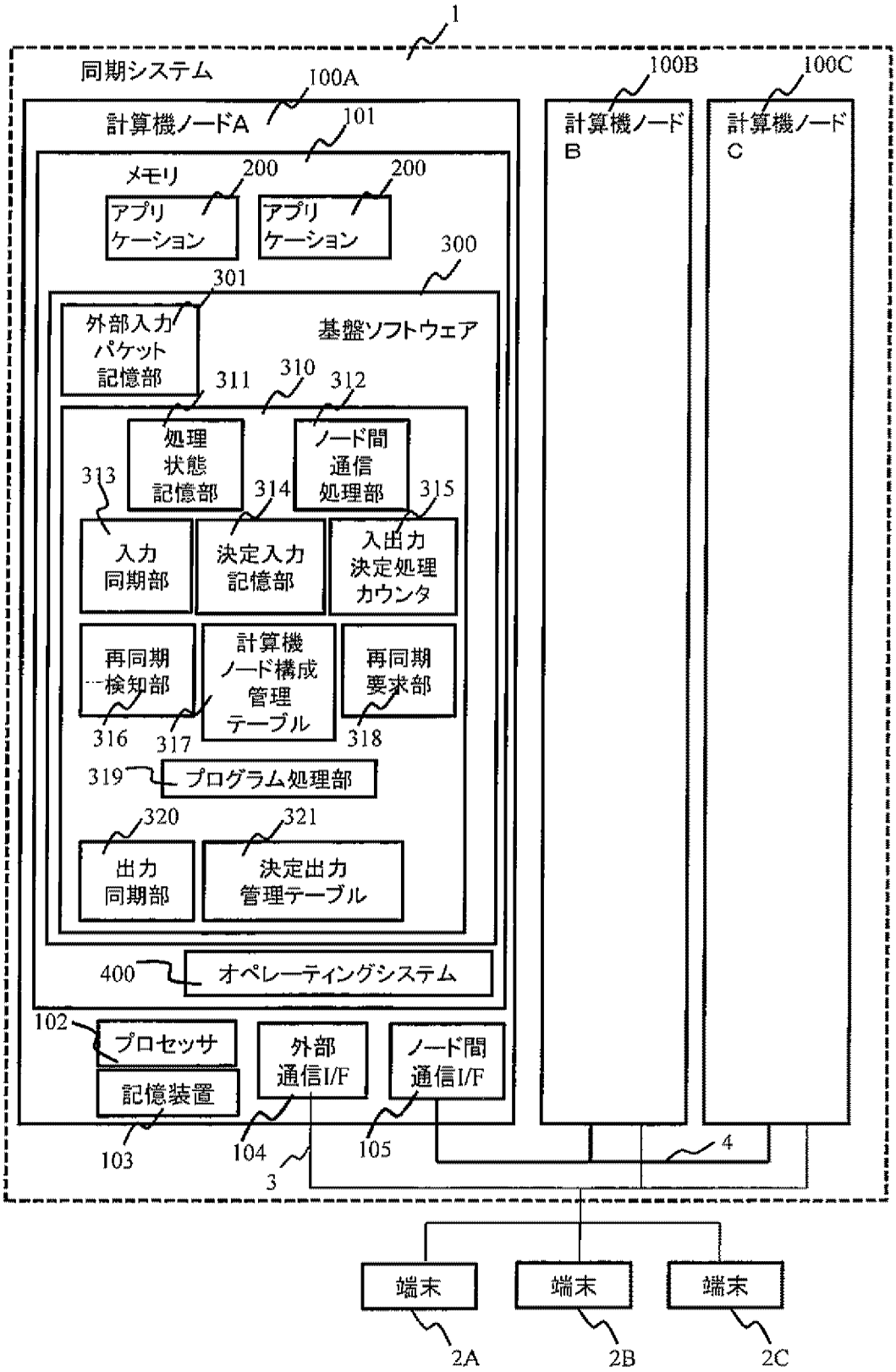
[請求項12] 請求項３記載の計算機システムであって、

前記各計算機ノードの前記同期部は、前記自計算機ノードと同期をとっている計算機ノードの数を管理しており、

前記各計算機ノードの前記同期部は、他の計算機ノードから前記自計算機ノードと同期をとるための同期要求を受信した場合に、当該同期部の処理の進捗状況を示す処理状態情報を当該他の計算機ノードに送信し、当該他の計算機ノードからの応答を受信すると、前記計算機ノードの数を更新することを特徴とする計算機システム。

[図1]

図1


1
同期システム
100A
計算機ノードA
101
メモリ
200
200
アプリケーション
アプリケーション
300
301
外部入力パケット記憶部
基盤ソフトウェア
311
310
312
処理状態記憶部
ノード間通信処理部
313
314
315
入力同期部
決定入力記憶部
入出力決定処理カウンタ
再同期検知部
計算機ノード構成管理テーブル
再同期要求部
316
317
318
319
プログラム処理部
320
321
出力同期部
決定出力管理テーブル
400
オペレーティングシステム
102
プロセッサ
記憶装置
外部通信I/F
ノード間通信I/F
103
104
105
3
4
100B
計算機ノードB
100C
計算機ノードC
端末
端末
端末
2A
2B
2C

[図2]

図2

外部入力パケットフォーマット 500

501 送信元情報	502 シーケンス番号	503 送信タイムスタップ	504 順列処理要否フラグ	505 順列処理ブロック数	506 順列処理開始番号	507 コンテンツ 508 アプリ識別子	509 パラメータ
2-C	2C-1	HHMMSS (例:081203)	1	10	2C-1	AP1	P1

[図3]

図3

外部入力パケット管理テーブル 550

551 番号	501 送信元情報	502 シーケンス番号	503 送信タイムスタップ	504 順列処理要否フラグ	505 順列処理ブロック数	506 順列処理開始番号	507 コンテンツ 508 アプリ識別子	509 パラメータ
1	2-C	2C-1	081203	1	10	2C-1	AP1	P1
≒	≒	≒	≒	≒	≒	≒	≒	
N	2-A	2A-1	091005	0	0	0	AP1	P1

[図4]

図4

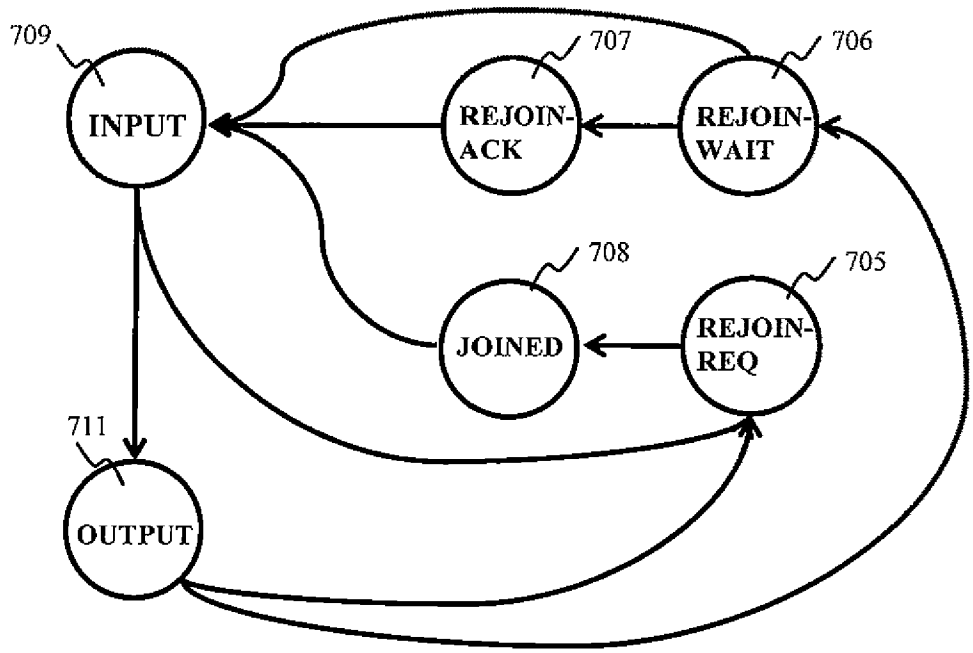
700

計算機ノード処理状態テーブル

701 処理状態識別子	702 付加情報	703 Penalty	704 処理状態
705 REJOIN-WAIT	—	0	REJOIN-WAIT
706 REJOIN-REQ	—	0	REJOIN-REQ
707 REJOIN-ACK	—	0	REJOIN-ACK
708 JOINED	—	0	JOINED
709 INPUT	710 入力同期ステップ Window Size	0	710-A BEGIN
			710-B END
711 OUTPUT	出力同期ステップ	—	—
712 OUTPUT		0	712-A BEGIN
			712-B END
713 LEADER		0	713-A BEGIN
			713-B END
714 VOTER		0	714-A LEADER
			714-B MONITOR

[図5]

図5



[図6]

図6

600

ノード間通信パケットフォーマット

601 計算機 ノード 識別子	602 入出力決定処理 カウンタ	603 計算機 ノード 処理状態	604 計算機 ノード 処理状態個別 詳細情報
		700	
100-A	C306	計算機ノード処理状態 テーブル内容	処理状態記憶部311 の結果

[図7]

図7

317

計算機ノード構成管理テーブル

317A

入力同期時計算機ノード総数		3
601		603
計算機ノード識別子	入出力決定処理カウンタ 602	計算機ノード処理状態
100A	C306	計算機ノード100A処理状態テーブル内容
100B	C306	計算機ノード100B処理状態テーブル内容
100C	C306	計算機ノード100C処理状態テーブル内容

[図8]

図8

800

決定入力管理テーブル

801

501 決定入力総数		502 503	504	505	3 506	507	
送信元情報	シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-C	2C-1	081203	1	10	2C-1	AP1	P1
2-A	2A-2	081215	0	0	0	AP1	P1
2-B	2B-1	081214	0	0	0	AP1	P1

[図9]

図9

810

決定済み入力管理テーブル

811

501 決定済み入力総数		503	504	505	3 506	507	
送信元情報	502 シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-B	2B-3	085234	0	0	0	AP1	P1

[図10]

図10

321

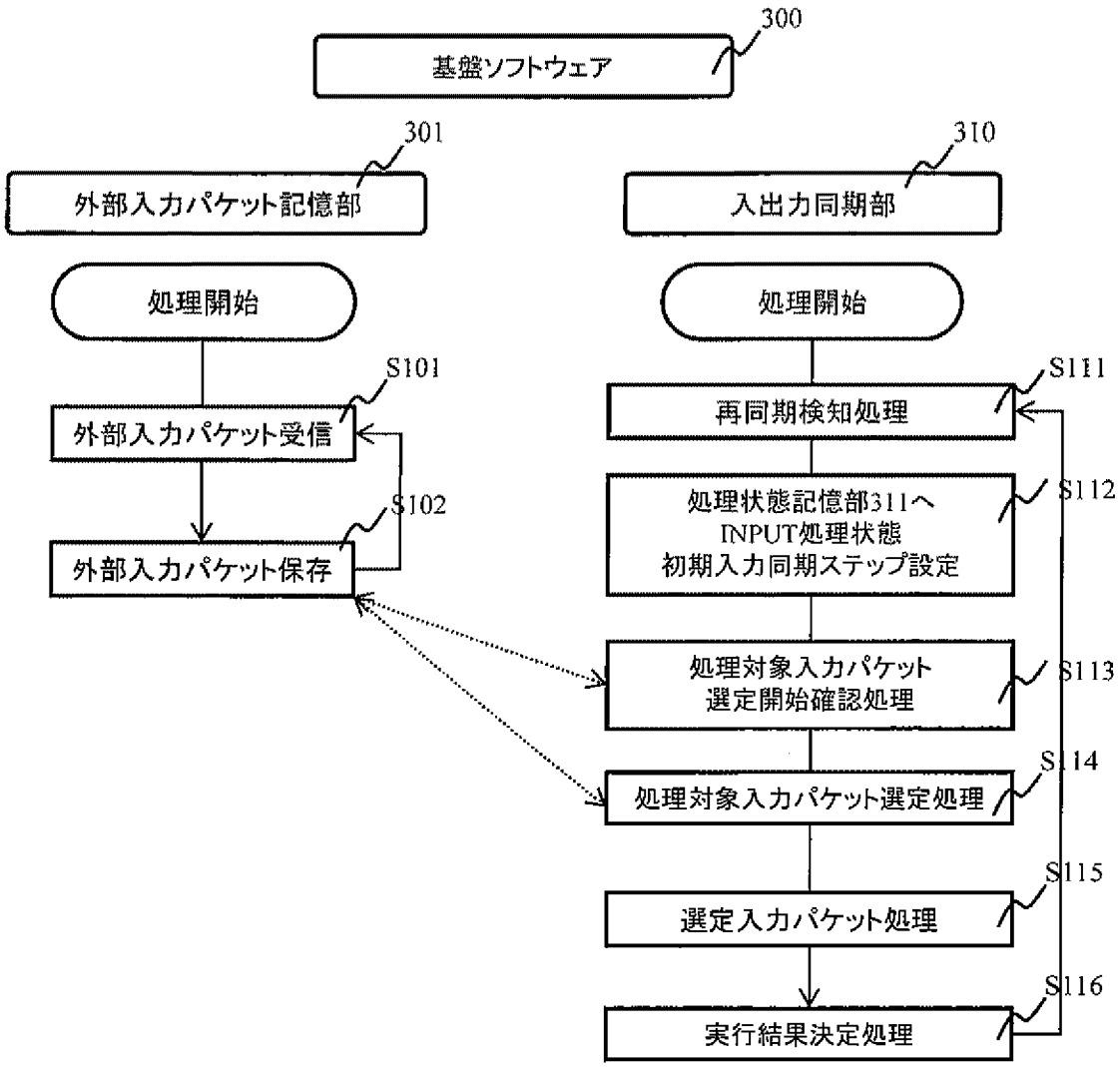
決定出力管理テーブル

321A

出力同期時計算機ノード総数			3	
順位	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態	計算機ノード処理結果
1	100C	C306	計算機ノード100C処理状態テーブル内容	X
2	100B	C306	計算機ノード100B処理状態テーブル内容	X
3	100A	C306	計算機ノード100A処理状態テーブル内容	X

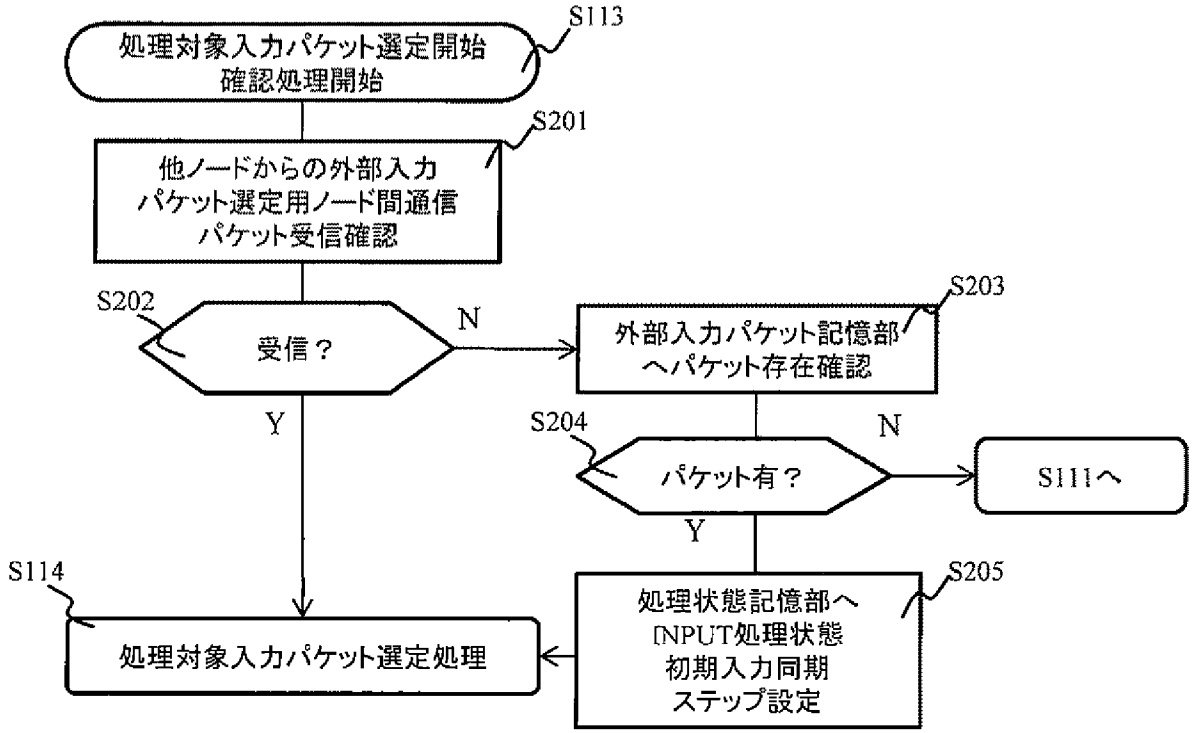
[図11]

図11


300
基盤ソフトウェア
301
外部入力パケット記憶部
310
入出力同期部
処理開始
処理開始
S101
外部入力パケット受信
S102
外部入力パケット保存
S111
再同期検知処理
S112
処理状態記憶部311へ
INPUT処理状態
初期入力同期ステップ設定
S113
処理対象入力パケット
選定開始確認処理
S114
処理対象入力パケット選定処理
S115
選定入力パケット処理
S116
実行結果決定処理

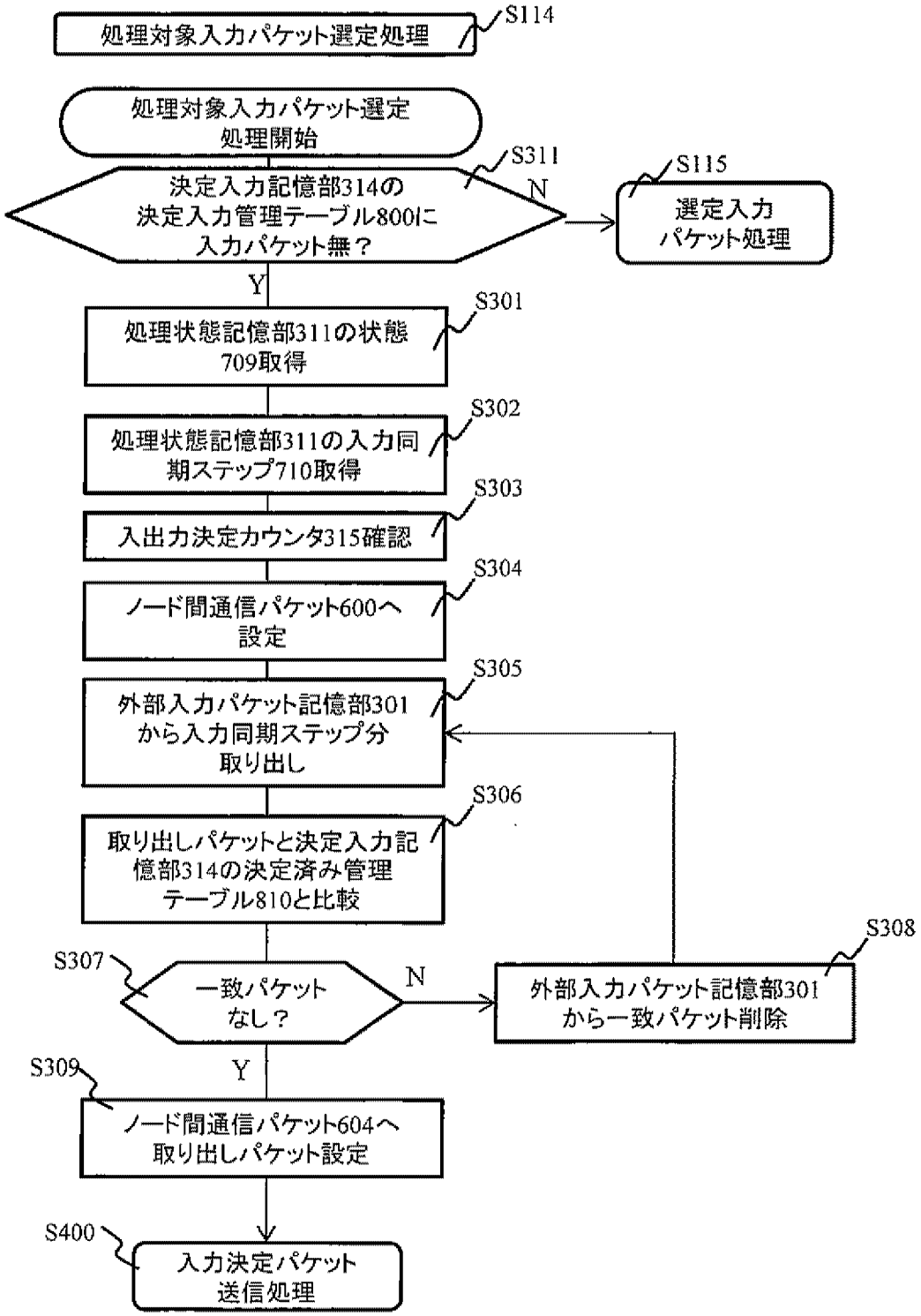
[図12]

図12


S113
処理対象入力パケット選定開始
確認処理開始
S201
他ノードからの外部入力
パケット選定用ノード間通信
パケット受信確認
S202
受信？
N
S203
外部入力パケット記憶部
へパケット存在確認
Y
S204
パケット有？
N
S111へ
Y
S114
処理対象入力パケット選定処理
S205
処理状態記憶部へ
INPUT処理状態
初期入力同期
ステップ設定

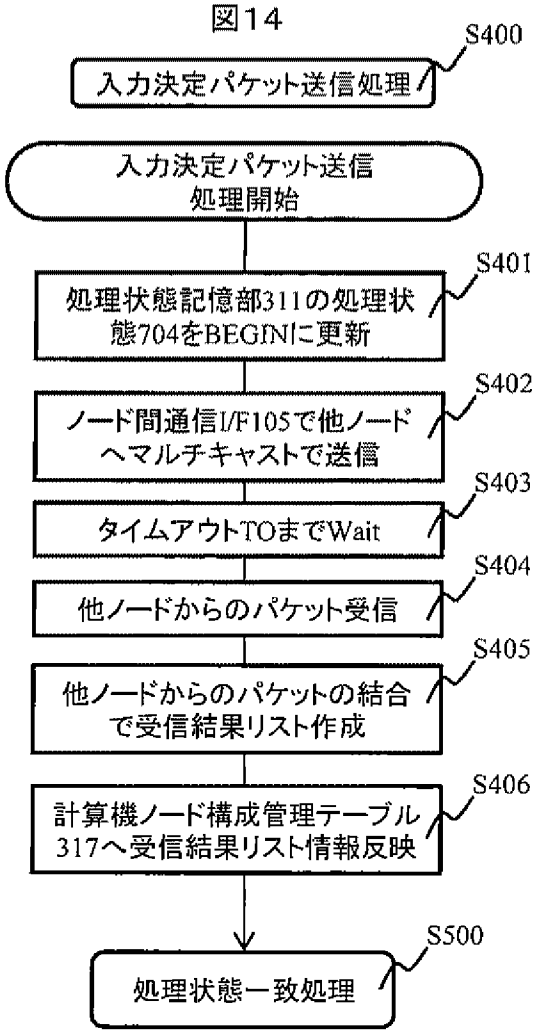
[図13]

図13


処理対象入力パケット選定処理
S114
処理対象入力パケット選定処理開始
S311
決定入力記憶部314の決定入力管理テーブル800に入力パケット無？
N
S115
選定入力パケット処理
Y
S301
処理状態記憶部311の状態709取得
S302
処理状態記憶部311の入力同期ステップ710取得
S303
入出力決定カウンタ315確認
S304
ノード間通信パケット600へ設定
S305
外部入力パケット記憶部301から入力同期ステップ分取り出し
S306
取り出しパケットと決定入力記憶部314の決定済み管理テーブル810と比較
S307
一致パケットなし？
N
S308
外部入力パケット記憶部301から一致パケット削除
Y
S309
ノード間通信パケット604へ取り出しパケット設定
S400
入力決定パケット送信処理

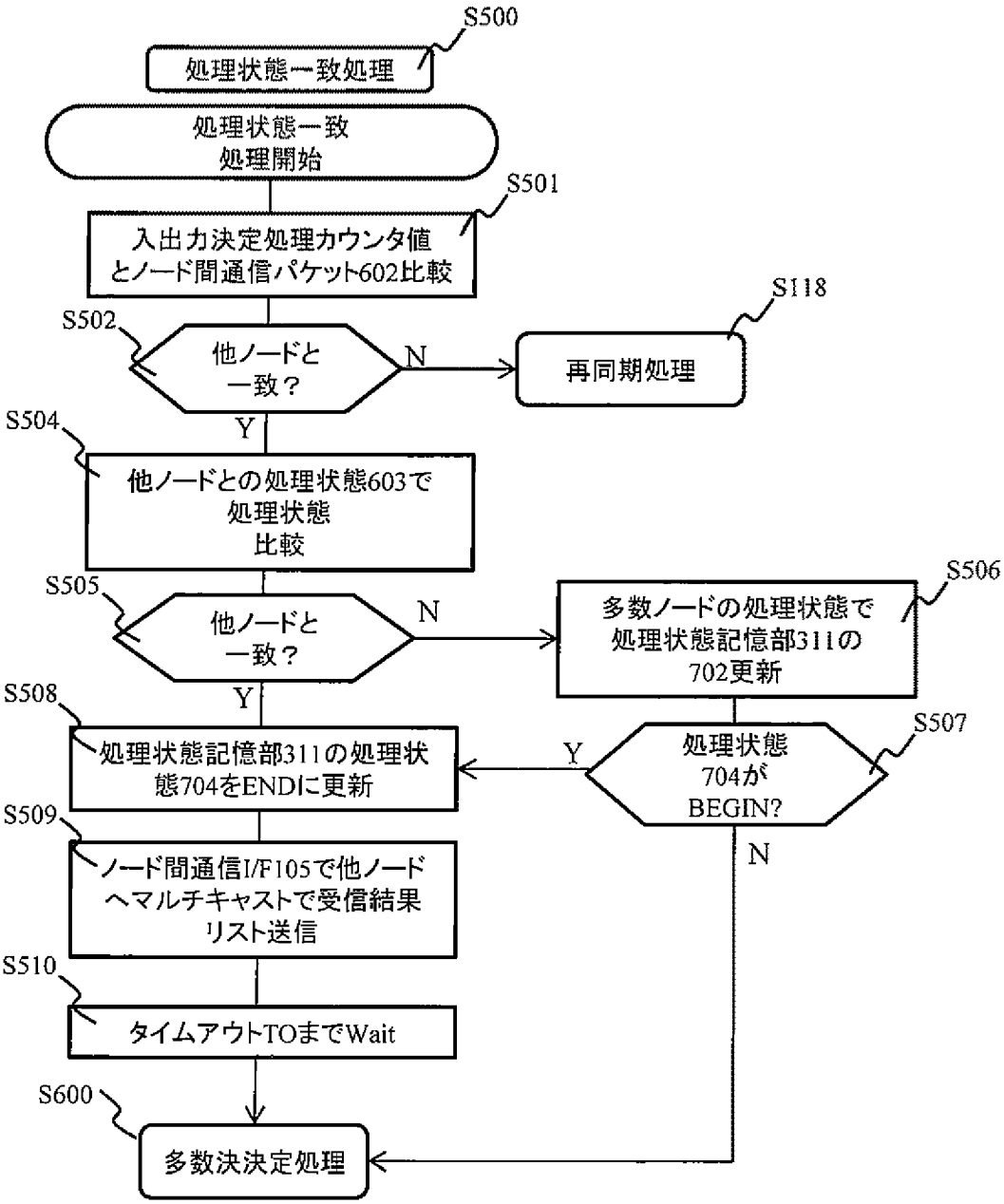
[図14]

図14


S400
入力決定パケット送信処理
入力決定パケット送信
処理開始
S401
処理状態記憶部311の処理状
態704をBEGINに更新
S402
ノード間通信I/F105で他ノード
へマルチキャストで送信
S403
タイムアウトTOまでWait
S404
他ノードからのパケット受信
S405
他ノードからのパケットの結合
で受信結果リスト作成
S406
計算機ノード構成管理テーブル
317へ受信結果リスト情報反映
S500
処理状態一致処理

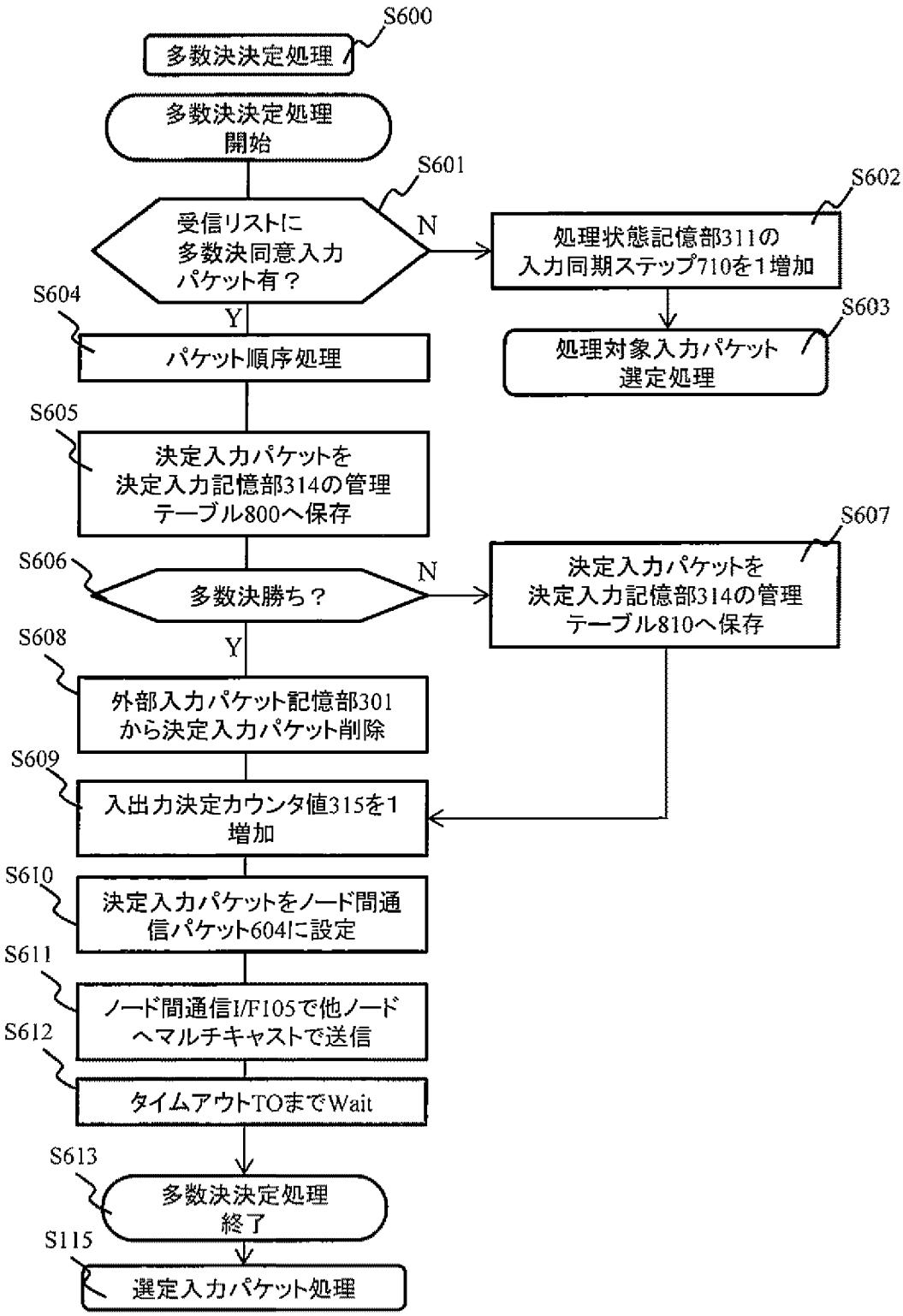
[図15]

図15


S500
処理状態一致処理
処理状態一致
処理開始
S501
入出力決定処理カウンタ値
とノード間通信パケット602比較
S502
他ノードと
一致？
N
S118
再同期処理
Y
S504
他ノードとの処理状態603で
処理状態
比較
S505
他ノードと
一致？
N
S506
多数ノードの処理状態で
処理状態記憶部311の
702更新
Y
S508
処理状態記憶部311の処理状
態704をENDに更新
Y
S507
処理状態
704が
BEGIN?
N
S509
ノード間通信I/F105で他ノード
へマルチキャストで受信結果
リスト送信
S510
タイムアウトTOまでWait
S600
多数決決定処理

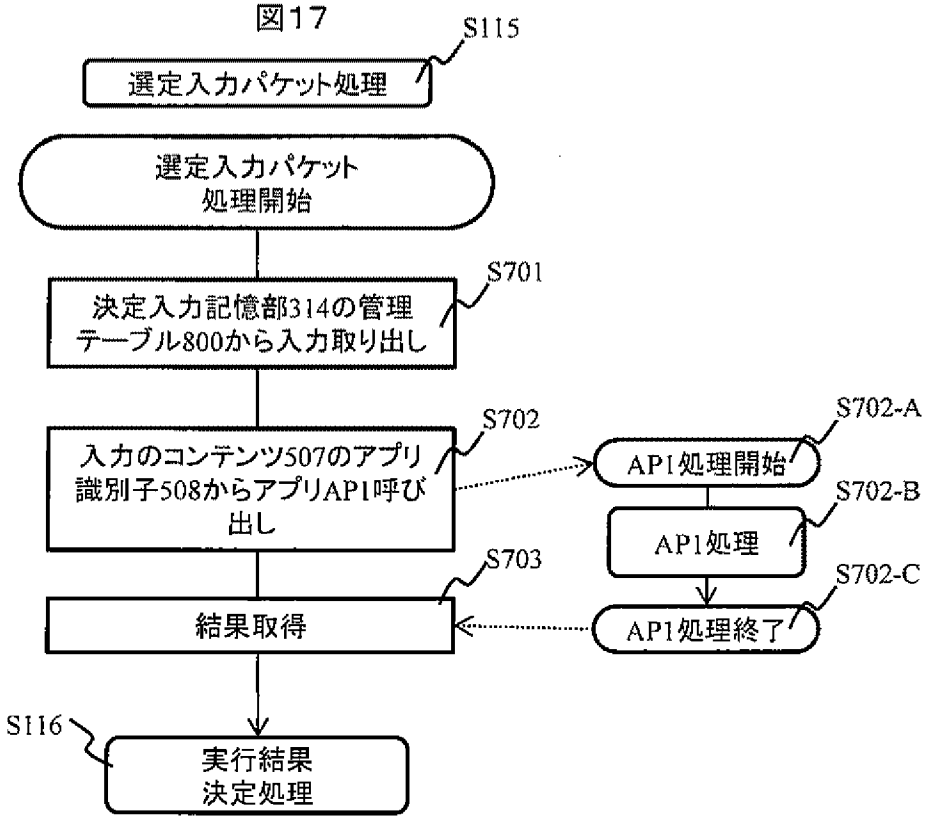
[図16]

図16


S600
多数決決定処理
多数決決定処理
開始
S601
受信リストに
多数決同意入力
パケット有？
N
S602
処理状態記憶部311の
入力同期ステップ710を1増加
S603
処理対象入力パケット
選定処理
Y
S604
パケット順序処理
S605
決定入力パケットを
決定入力記憶部314の管理
テーブル800へ保存
S606
多数決勝ち？
N
S607
決定入力パケットを
決定入力記憶部314の管理
テーブル810へ保存
Y
S608
外部入力パケット記憶部301
から決定入力パケット削除
S609
入出力決定カウンタ値315を1
増加
S610
決定入力パケットをノード間通
信パケット604に設定
S611
ノード間通信I/F105で他ノード
へマルチキャストで送信
S612
タイムアウトTOまでWait
S613
多数決決定処理
終了
S115
選定入力パケット処理

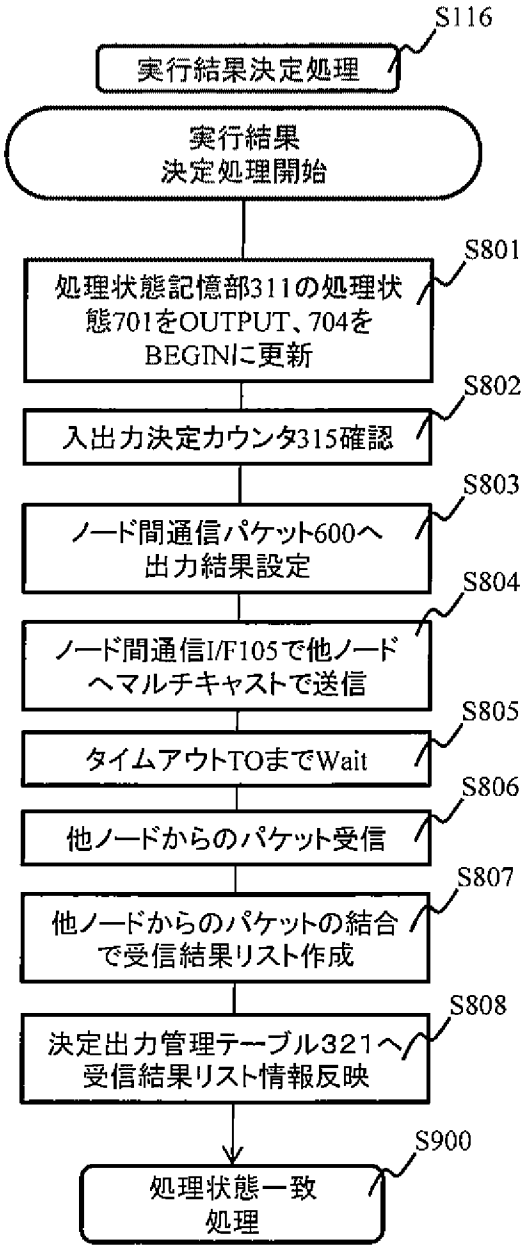
[図17]

図17


S115
選定入力パケット処理
選定入力パケット
処理開始
S701
決定入力記憶部314の管理
テーブル800から入力取り出し
S702
入力のコンテンツ507のアプリ
識別子508からアプリAPI呼び
出し
S702-A
API処理開始
S702-B
API処理
S702-C
API処理終了
S703
結果取得
S116
実行結果
決定処理

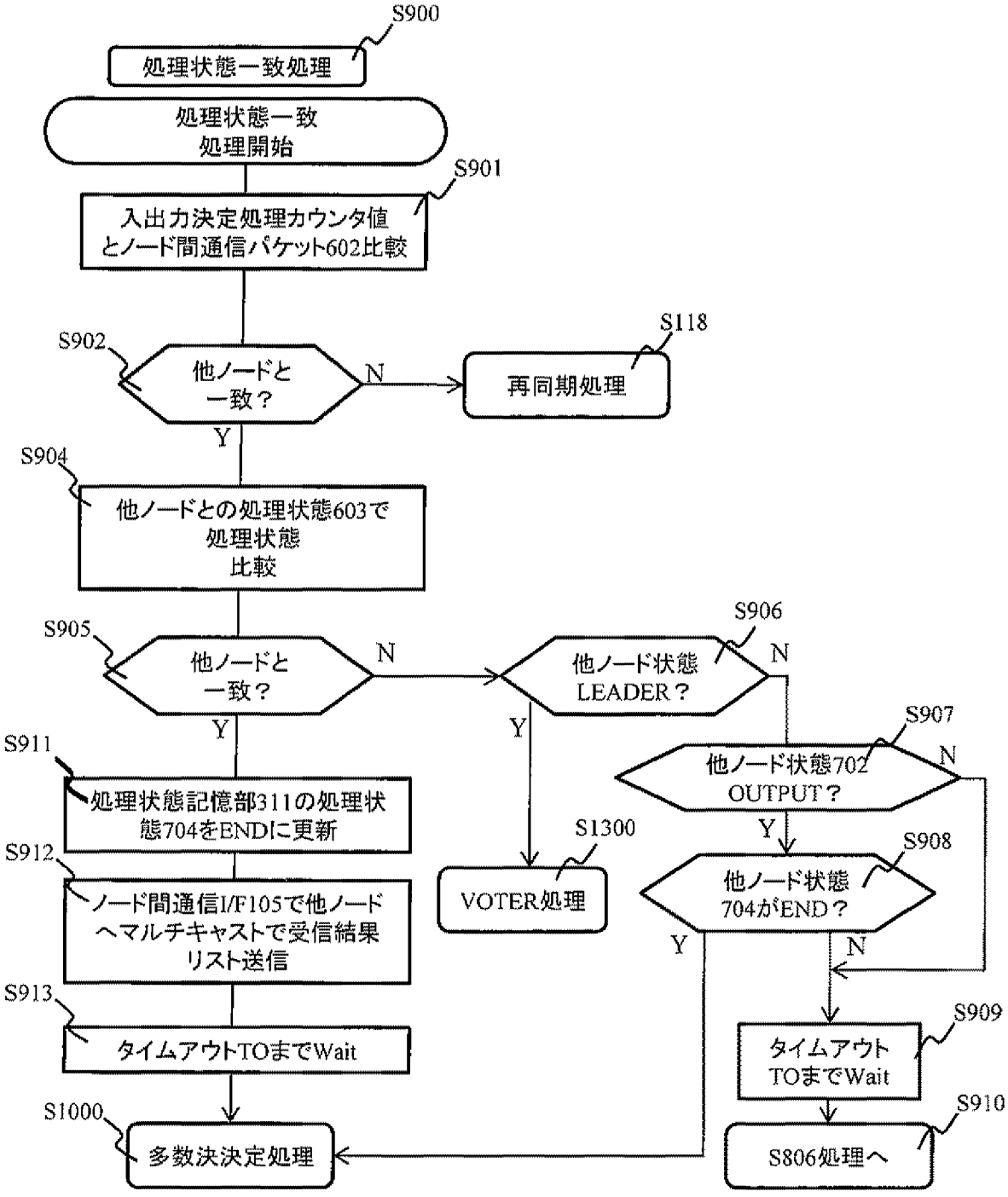
[図18]

図18


S116
実行結果決定処理
実行結果
決定処理開始
S801
処理状態記憶部311の処理状態701をOUTPUT、704をBEGINに更新
S802
入出力決定カウンタ315確認
S803
ノード間通信パケット600へ出力結果設定
S804
ノード間通信I/F105で他ノードへマルチキャストで送信
S805
タイムアウトTOまでWait
S806
他ノードからのパケット受信
S807
他ノードからのパケットの結合で受信結果リスト作成
S808
決定出力管理テーブル321へ受信結果リスト情報反映
S900
処理状態一致処理

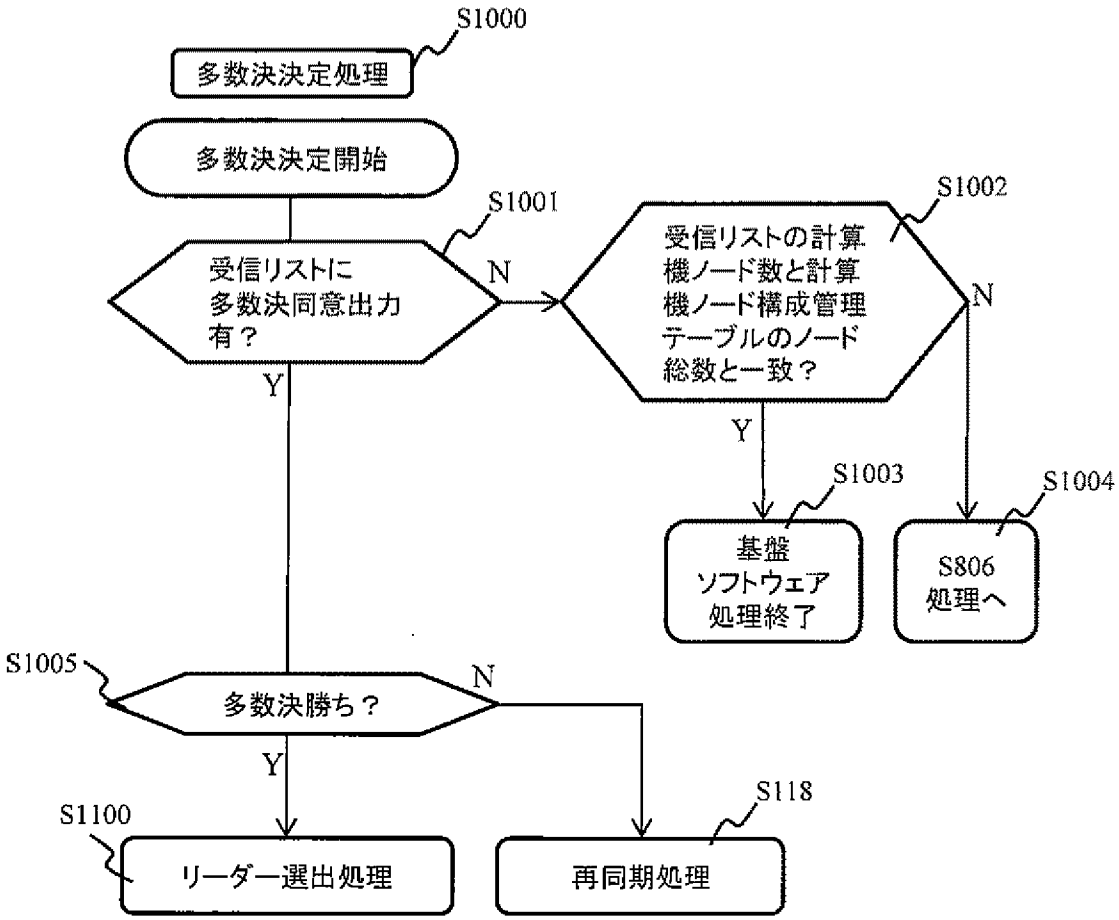
[図19]

図19


S900
処理状態一致処理
処理状態一致
処理開始
S901
入出力決定処理カウンタ値
とノード間通信パケット602比較
S902
他ノードと
一致？
N
S118
再同期処理
Y
S904
他ノードとの処理状態603で
処理状態
比較
S905
他ノードと
一致？
N
S906
他ノード状態
LEADER？
N
Y
Y
S907
他ノード状態702
OUTPUT？
N
S911
処理状態記憶部311の処理状
態704をENDに更新
Y
S1300
VOTER処理
S908
他ノード状態
704がEND？
S912
ノード間通信I/F105で他ノード
へマルチキャストで受信結果
リスト送信
Y
N
S913
タイムアウトTOまでWait
S909
タイムアウト
TOまでWait
S1000
多数決決定処理
S910
S806処理へ

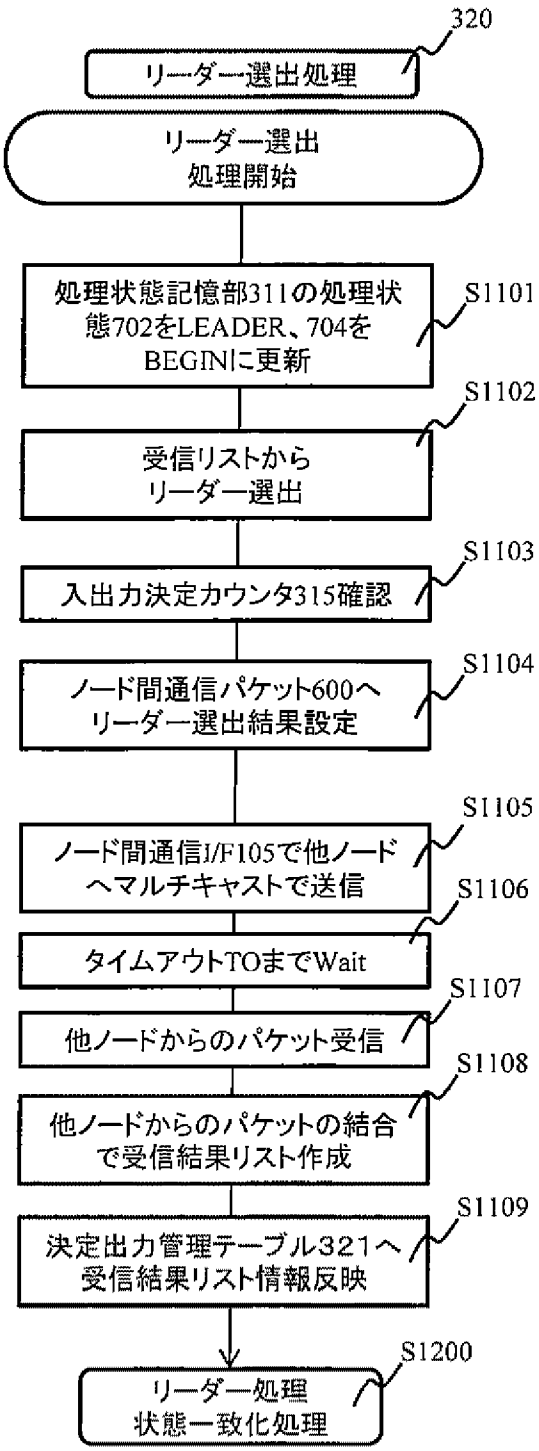
[図20]

図20


S1000
多数決決定処理
多数決決定開始
S1001
受信リストに
多数決同意出力
有？
N
S1002
受信リストの計算
機ノード数と計算
機ノード構成管理
テーブルのノード
総数と一致？
N
Y
Y
S1003
基盤
ソフトウェア
処理終了
S1004
S806
処理へ
S1005
多数決勝ち？
N
Y
S1100
リーダー選出処理
S118
再同期処理

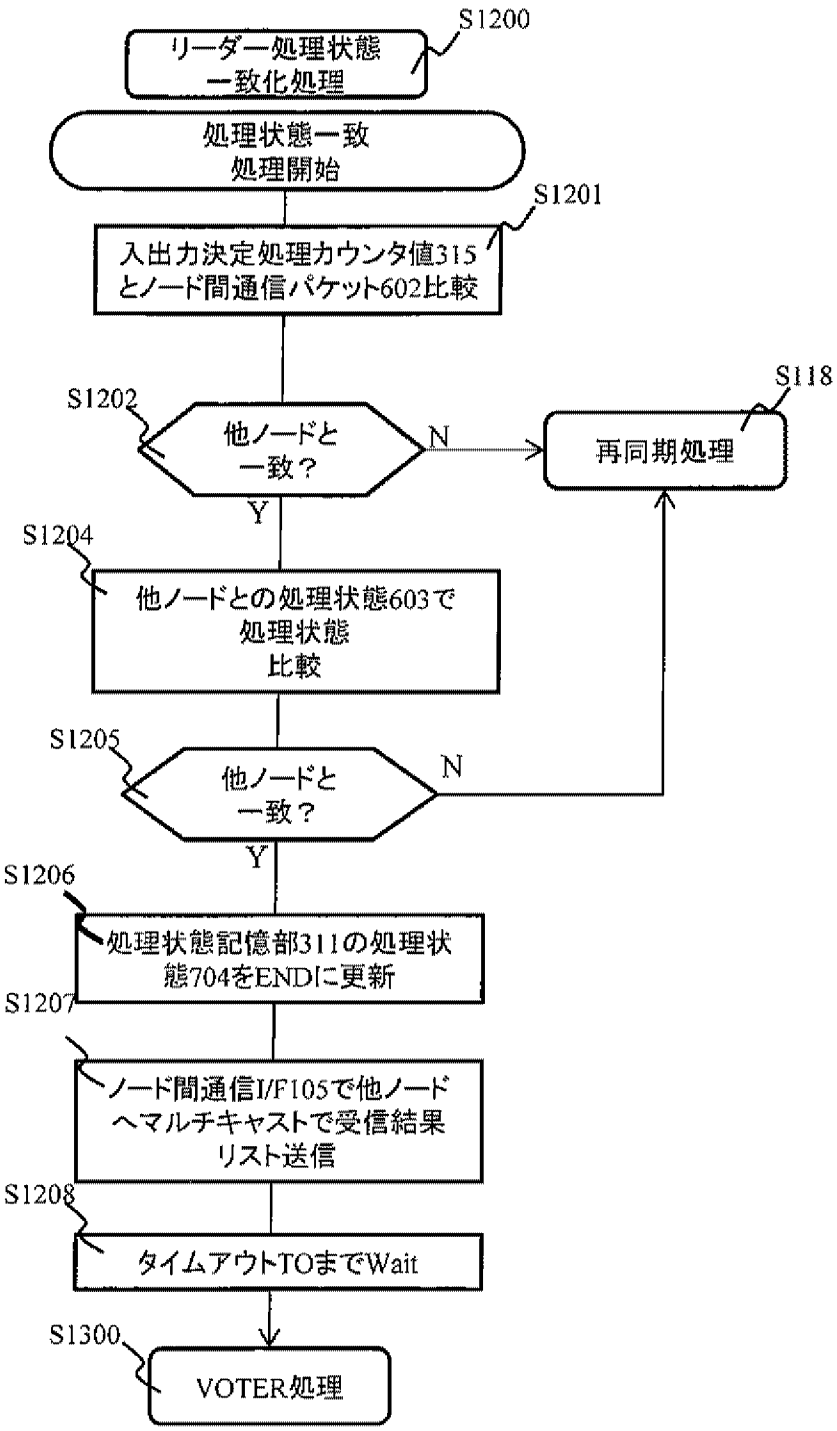
[図21]

図21


320
リーダー選出処理
リーダー選出
処理開始
処理状態記憶部311の処理状態702をLEADER、704をBEGINに更新
S1101
S1102
受信リストから
リーダー選出
S1103
入出力決定カウンタ315確認
S1104
ノード間通信パケット600へ
リーダー選出結果設定
S1105
ノード間通信I/F105で他ノードへマルチキャストで送信
S1106
タイムアウトTOまでWait
S1107
他ノードからのパケット受信
S1108
他ノードからのパケットの結合で受信結果リスト作成
S1109
決定出力管理テーブル321へ
受信結果リスト情報反映
S1200
リーダー処理
状態一致化処理

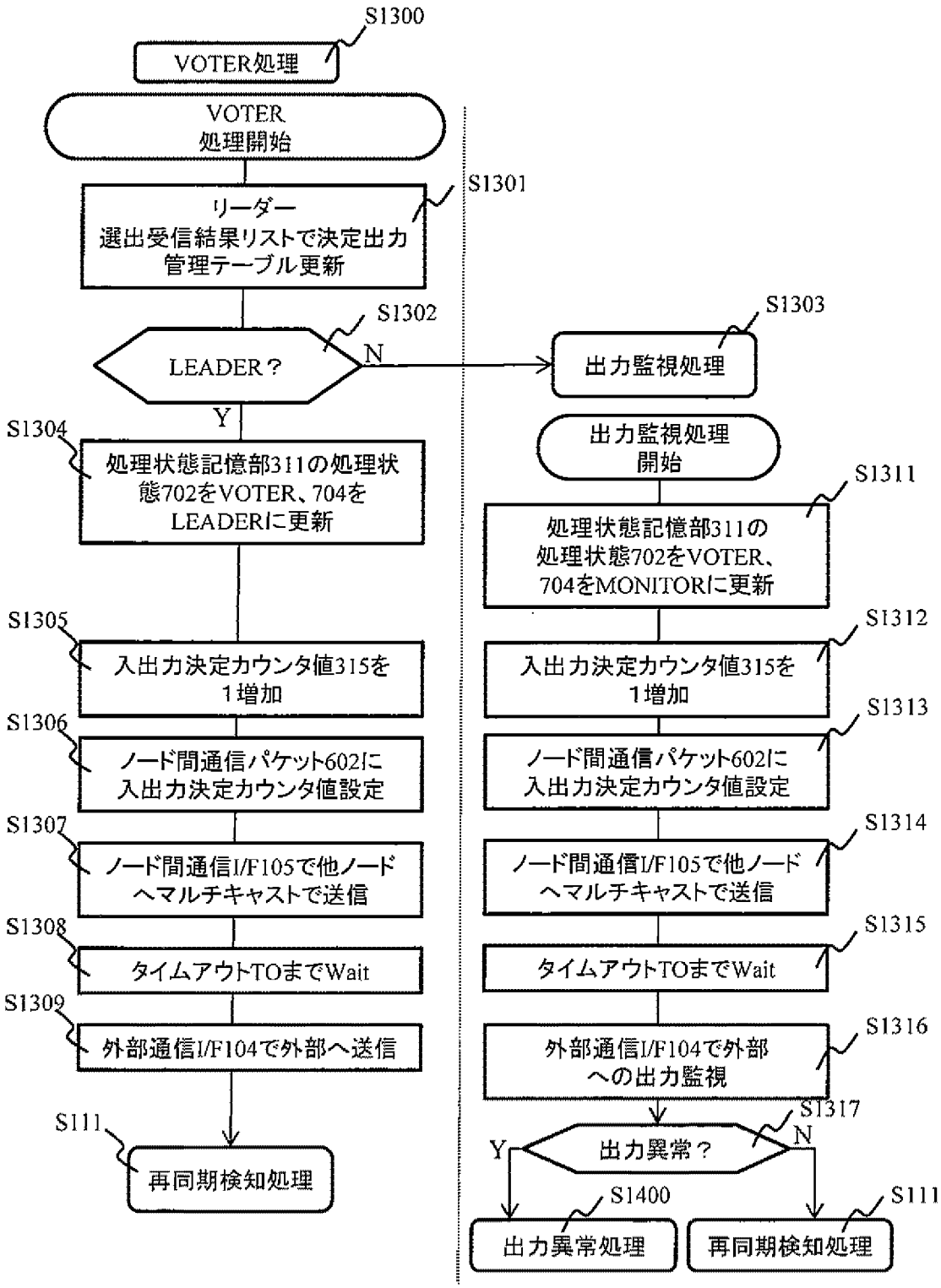
[図22]

図22


S1200
リーダー処理状態
一致化処理
処理状態一致
処理開始
S1201
入出力決定処理カウンタ値315
とノード間通信パケット602比較
S1202
他ノードと
一致？
N
S118
再同期処理
Y
S1204
他ノードとの処理状態603で
処理状態
比較
S1205
他ノードと
一致？
N
Y
S1206
処理状態記憶部311の処理状
態704をENDに更新
S1207
ノード間通信I/F105で他ノード
へマルチキャストで受信結果
リスト送信
S1208
タイムアウトTOまでWait
S1300
VOTER処理

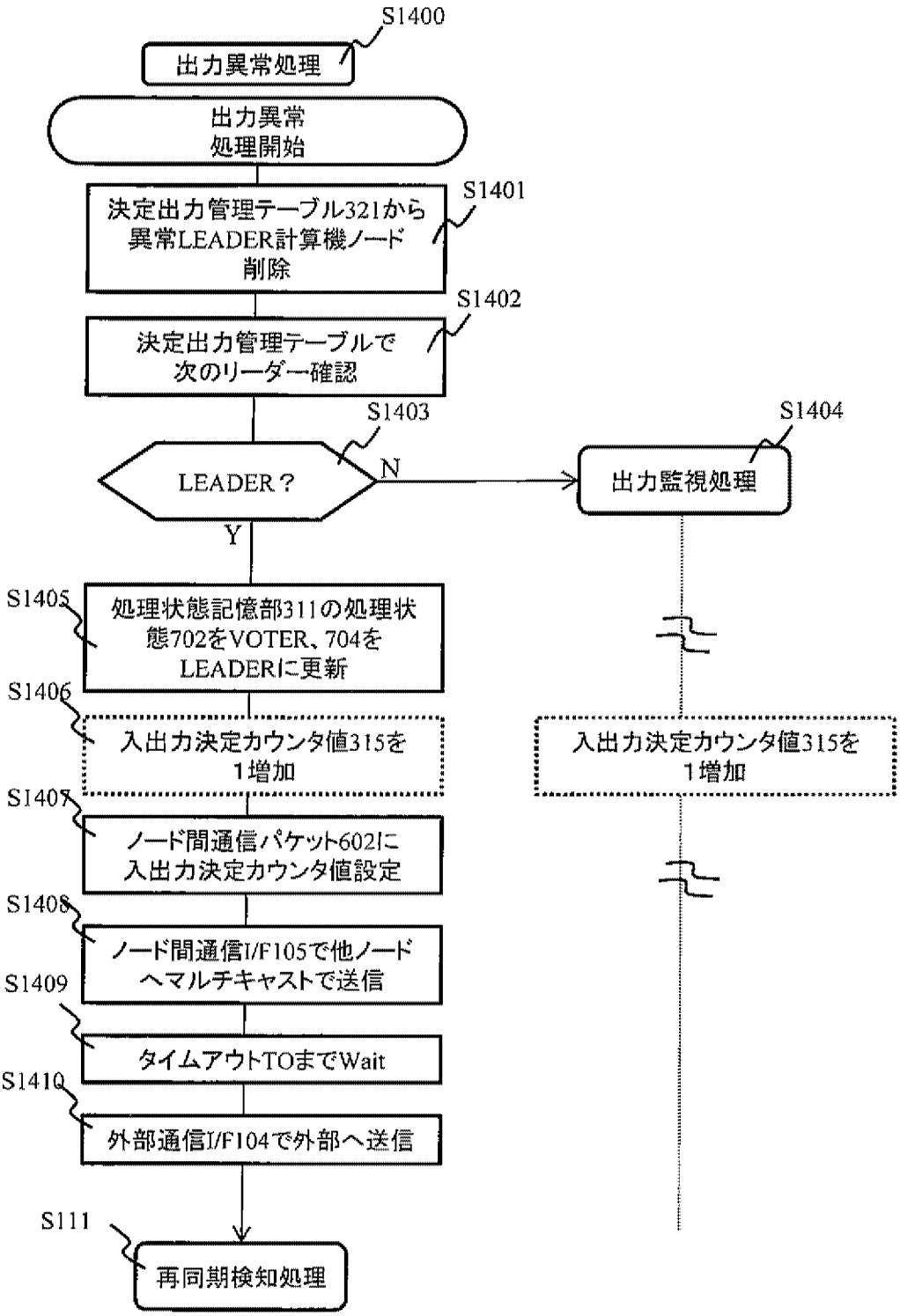
[図23]

図23


S1300
VOTER処理
VOTER
処理開始
S1301
リーダー
選出受信結果リストで決定出力
管理テーブル更新
S1302
LEADER？
N
Y
S1303
出力監視処理
S1304
処理状態記憶部311の処理状
態702をVOTER、704を
LEADERに更新
出力監視処理
開始
S1311
処理状態記憶部311の
処理状態702をVOTER、
704をMONITORに更新
S1305
入出力決定カウンタ値315を
1増加
S1312
入出力決定カウンタ値315を
1増加
S1306
ノード間通信パケット602に
入出力決定カウンタ値設定
S1313
ノード間通信パケット602に
入出力決定カウンタ値設定
S1307
ノード間通信I/F105で他ノード
へマルチキャストで送信
S1314
ノード間通信I/F105で他ノード
へマルチキャストで送信
S1308
タイムアウトTOまでWait
S1315
タイムアウトTOまでWait
S1309
外部通信I/F104で外部へ送信
S1316
外部通信I/F104で外部
への出力監視
S1317
出力異常？
Y
N
S111
再同期検知処理
S1400
出力異常処理
S111
再同期検知処理

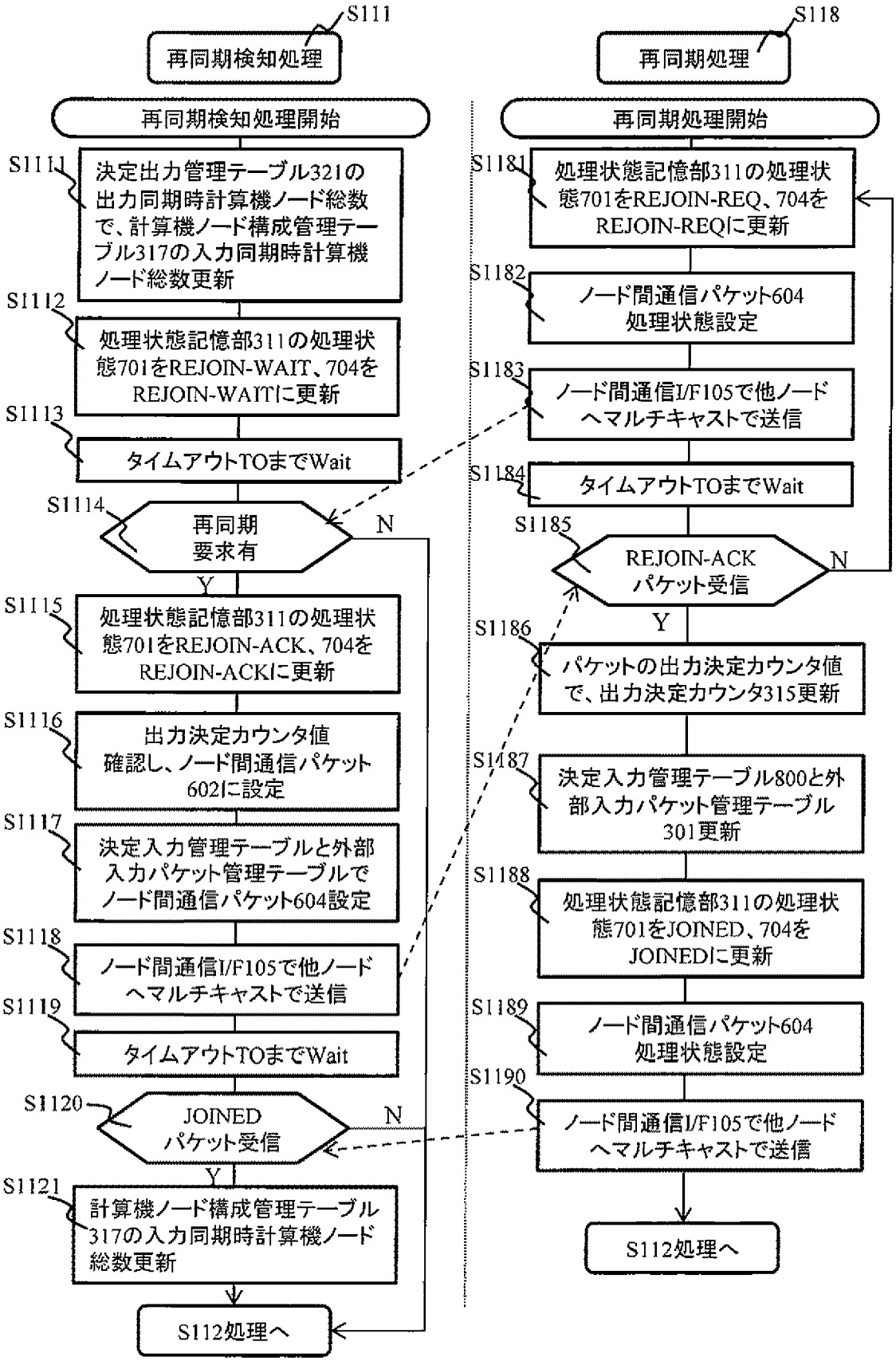
[図24]

図24


S1400
出力異常処理
出力異常
処理開始
S1401
決定出力管理テーブル321から
異常LEADER計算機ノード
削除
S1402
決定出力管理テーブルで
次のリーダー確認
S1403
LEADER？
N
S1404
出力監視処理
Y
S1405
処理状態記憶部311の処理状
態702をVOTER、704を
LEADERに更新
S1406
入出力決定カウンタ値315を
1増加
入出力決定カウンタ値315を
1増加
S1407
ノード間通信パケット602に
入出力決定カウンタ値設定
S1408
ノード間通信I/F105で他ノード
へマルチキャストで送信
S1409
タイムアウトTOまでWait
S1410
外部通信I/F104で外部へ送信
S111
再同期検知処理

[図25]

図25


S111
再同期検知処理
再同期検知処理開始
S1111
決定出力管理テーブル321の出力同期時計算機ノード総数で、計算機ノード構成管理テーブル317の入力同期時計算機ノード総数更新
S1112
処理状態記憶部311の処理状態701をREJOIN-WAIT、704をREJOIN-WAITに更新
S1113
タイムアウトTOまでWait
S1114
再同期要求有
N
Y
S1115
処理状態記憶部311の処理状態701をREJOIN-ACK、704をREJOIN-ACKに更新
S1116
出力決定カウンタ値確認し、ノード間通信パケット602に設定
S1117
決定入力管理テーブルと外部入力パケット管理テーブルでノード間通信パケット604設定
S1118
ノード間通信I/F105で他ノードへマルチキャストで送信
S1119
タイムアウトTOまでWait
S1120
JOINEDパケット受信
N
Y
S1121
計算機ノード構成管理テーブル317の入力同期時計算機ノード総数更新
S112処理へ
S118
再同期処理
再同期処理開始
S1181
処理状態記憶部311の処理状態701をREJOIN-REQ、704をREJOIN-REQに更新
S1182
ノード間通信パケット604処理状態設定
S1183
ノード間通信I/F105で他ノードへマルチキャストで送信
S1184
タイムアウトTOまでWait
S1185
REJOIN-ACKパケット受信
N
Y
S1186
パケットの出力決定カウンタ値で、出力決定カウンタ315更新
S1187
決定入力管理テーブル800と外部入力パケット管理テーブル301更新
S1188
処理状態記憶部311の処理状態701をJOINED、704をJOINEDに更新
S1189
ノード間通信パケット604処理状態設定
S1190
ノード間通信I/F105で他ノードへマルチキャストで送信
S112処理へ

[図26]

図26

(a)

REJOIN-REQのノード間通信パケット内容

計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				詳細情報
100A	0	REJOIN-REQ	0	0	REJOIN-REQ	−

(b)

REJOIN-ACKのノード間通信パケット内容

計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				詳細情報	
100B	C012	REJOIN-ACK	0	0	REJOIN-ACK	800	301

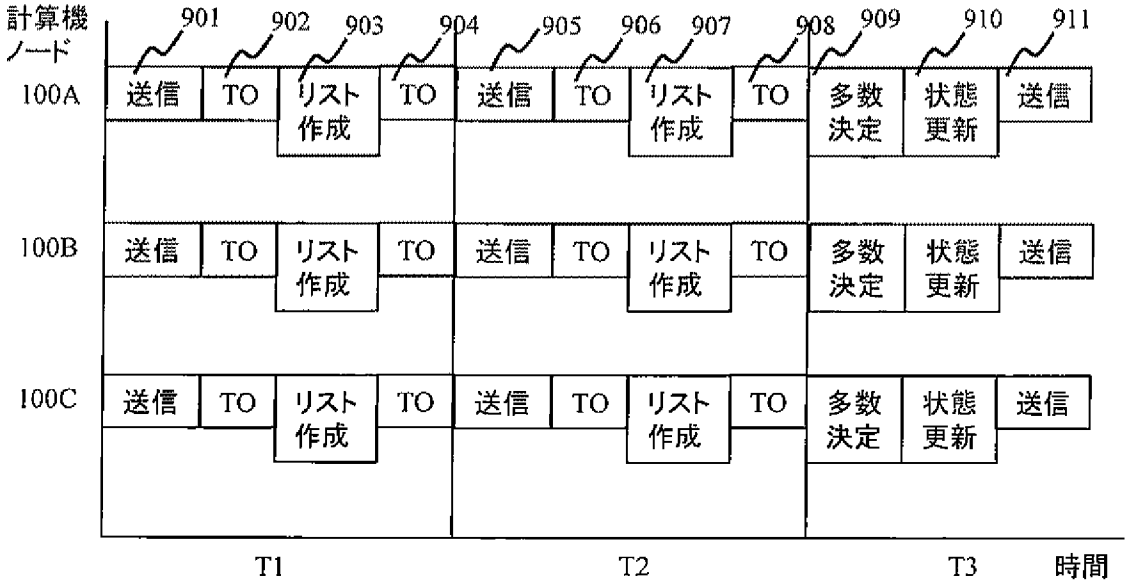
(c)

JOINEDのノード間通信パケット内容

計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				詳細情報
100A	C012	JOINED	0	0	JOINED	−

[図27]

図27



[図28]

図28

各時間Tに対する各計算機ノードの外部入力パケット記憶部、処理状態記憶部及び入力決定カウンタの状態

		計算機ノード100A				計算機ノード100B				計算機ノード100C			
	時間	SEQ	WS	ST	CNT	SEQ	WS	ST	CNT	SEQ	WS	ST	CNT
920	T1	2A-1	2	B	C010	0	2	B	C010	2C-1	2	B	C010
		2B-1				0				0			
		0				0				0			
921	T2	2A-1	3	B	C010	2B-1	3	B	C010	2C-1	3	B	C010
		2B-1				0				2B-1			
		2C-1				0				0			
922	T3	2A-1	3	E	C011	0	3	E	C011	0	3	E	C011
		0				0				0			
		0				0				0			

SEQ:外部入力パケット500のシーケンス番号
WS:処理状態テーブル700の入力同期ステップ
ST:処理状態テーブル700の処理状態
B:処理状態テーブル700のBEGIN状態
E:処理状態テーブル700のEND状態
CNT:入出力決定処理カウンタ

[図29]

図29

時間別送信に対する各計算機ノードのノード間通信パケット内容

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				計算機ノード処理状態個別詳細情報
				701	702	703	704	外部入力パケット×入力同期ステップ分
940	T1	100A	C010	INPUT	2	0	BEGIN	2A-I 2B-I
		100B	C010	INPUT	2	0	BEGIN	0 0
		100C	C010	INPUT	2	0	BEGIN	2C-I 0
941	T2	100A	C010	INPUT	3	0	BEGIN	2A-I 2B-I 2C-I
		100B	C010	INPUT	3	0	BEGIN	2B-I 0 0
		100C	C010	INPUT	3	0	BEGIN	2C-I 2B-I 0
942	T3	100A	C010	INPUT	3	0	END	2B-I 2C-I 0
		100B	C010	INPUT	3	0	END	2B-I 2C-I 0
		100C	C010	INPUT	3	0	END	2B-I 2C-I 0

936 943 944 945 937

[図30]

図30

時間別リスト作成に対する各計算機ノードのノード間通信パケット内容

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				計算機ノード処理状態個別詳細情報
				701	702	703	704	外部入力パケット×入力同期ステップ分
960	T1	100A	C010	INPUT	2	0	BEGIN	2A-I 2B-I
		100B	C010	INPUT	2	0	BEGIN	0 0
		100C	C010	INPUT	2	0	BEGIN	2C-I 0
961	T2	100A	C010	INPUT	3	0	BEGIN	2A-I 2B-I 2C-I
		100B	C010	INPUT	3	0	BEGIN	2B-I 0 0
		100C	C010	INPUT	3	0	BEGIN	2C-I 2B-I 0

965 962 963 964 966

[図31]

図31

時間別計算機ノード構成管理テーブルの変化

	時間	入力同期時計算機ノード総数		3			
		計算機ノード識別子	入出力決定処理カウンタ			計算機ノード処理状態	
				処理状態識別子	付加情報	Penalty	処理状態
980	T1	100A	C010	INPUT	2	0	E
		100B	C010	INPUT	2	0	E
		100C	C010	INPUT	2	0	E
981	T2	100A	C010	INPUT	3	0	E
		100B	C010	INPUT	3	0	E
		100C	C010	INPUT	3	0	E
982	T3	100A	C011	INPUT	3	0	E
		100B	C011	INPUT	3	0	E
		100C	C011	INPUT	3	0	E

983 984

[図32]

図32

計算機ノード100A、100B、100Cの決定入力管理テーブル

決定入力総数			2				
送信元情報	シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-B	2B-1	081214	0	0	0	AP1	P1
2-C	2C-1	081215	0	0	0	AP1	P1

[図33]

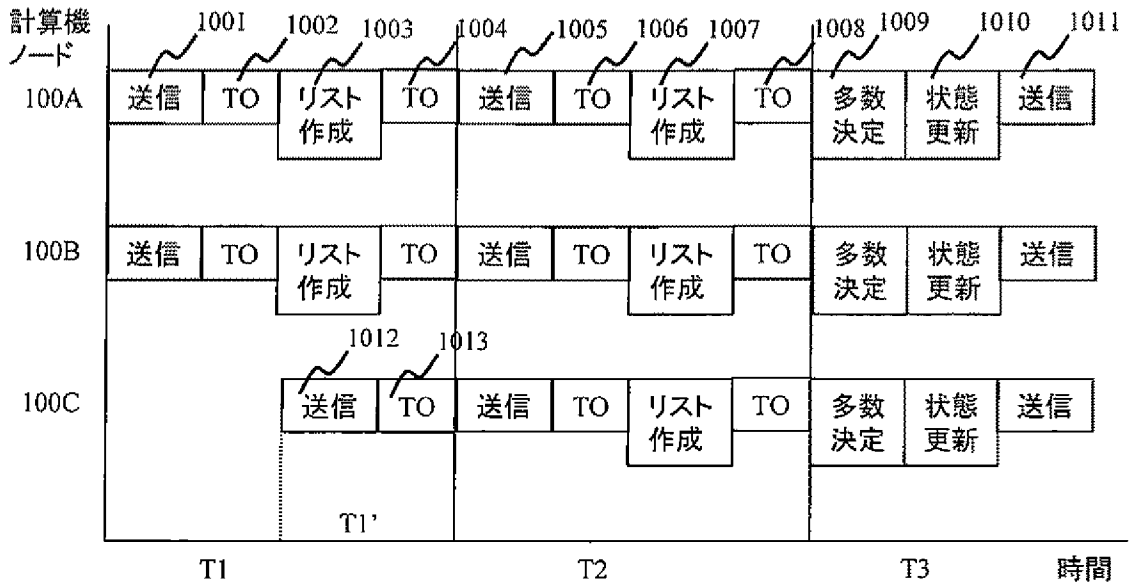
図33

計算機ノード100Bの決定済み入力管理テーブル

決定済み入力総数			1				
送信元情報	シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-B	2B-1	081214	0	0	0	AP1	P1

[図34]

図34



[図35]

図35

		計算機ノード100A				計算機ノード100B				計算機ノード100C			
	時間	SEQ	WS	ST	CNT	SEQ	WS	ST	CNT	SEQ	WS	ST	CNT
1028	T1	2A-1	2	B	C010	2C-1	2	B	C010	0	2	B	C010
		2B-1	1031			0	1032			0	1033		
		0				0				0			
1029	T2	2A-1	3	B	C010	2C-1	3	B	C010	2A-1	3	B	C010
		2B-1				0				0	1034		
		0				0				0			
1030	T3	2B-1	3	E	C011	2C-1	3	E	C011	0	3	E	C011
		0				0				0			
		0				0				0			

SEQ: 外部入力パケット500のシーケンス番号
WS: 処理状態テーブル700の入力同期ステップ
ST: 処理状態テーブル700の処理状態
B: 処理状態テーブル700のBEGIN状態
E: 処理状態テーブル700のEND状態
CNT: 入出力決定処理カウンタ

[図36]

図36

時間別送信に対する各計算機ノードのノード間通信パケット内容

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				計算機ノード処理状態個別詳細情報
				701	702	703	704	外部入力パケット×入力同期ステップ分
1040	T1	100A	C010	INPUT	2	0	BEGIN	2A-1 2B-1
1041		100B	C010	INPUT	2	0	BEGIN	2C-1 0
	T1'	100C	C010	INPUT	2	0	BEGIN	2A-1 0
1042	T2	100A	C010	INPUT	3	0	BEGIN	2A-1 2B-1 0
		100B	C010	INPUT	3	0	BEGIN	2C-1 0 0
		100C	C010	INPUT	3	1	BEGIN	2A-1 0 0
1043	T3	100A	C010	INPUT	3	0	END	2A-1 0 0
		100B	C010	INPUT	3	0	END	2A-1 0 0
		100C	C010	INPUT	3	1	END	2A-1 0 0

1044 1045 1046 1047

[図37]

図37

時間別リスト作成に対する各計算機ノードのノード間通信パケット内容

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				計算機ノード処理状態個別詳細情報
				701	702	703	704	外部入力パケット×入力同期ステップ分
1060	T1	100A	C010	INPUT	2	0	BEGIN	2A-1 2B-1
		100B	C010	INPUT	2	0	BEGIN	2C-1 0
1061	T2	100A	C010	INPUT	3	0	BEGIN	2A-1 2B-1 0
		100B	C010	INPUT	3	0	BEGIN	2C-1 0 0
		100C	C010	INPUT	3	1	BEGIN	2A-1 0 0

1062 1063 1064 1065 1066

[図38]

図38

時間別計算機ノード100A、100B構成管理テーブル変化

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態			
				処理状態識別子	付加情報	Penalty	処理状態
1080	T1	100A	C010	INPUT	2	0	E
		100B	C010	INPUT	2	0	E
1081	T2	100A	C010	INPUT	3	0	E
		100B	C010	INPUT	3	0	E
		100C	C010	INPUT	3	1	E
1082	T3	100A	C011	INPUT	3	0	E
		100B	C011	INPUT	3	0	E
		100C	C011	INPUT	3	1	E

1083 1084 1085 1086 1087

[図39]

図39

時間別計算機ノード100C構成管理テーブル変化

	時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態			
				処理状態識別子	付加情報	Penalty	処理状態
1090	T1'						
		100C	C010	INPUT	2	0	E
1091	T2	100A	C010	INPUT	3	0	E
		100B	C010	INPUT	3	0	E
		100C	C010	INPUT	3	1	E
1092	T3	100A	C011	INPUT	3	0	E
		100B	C011	INPUT	3	0	E
		100C	C011	INPUT	3	1	E

1093 1094 1095 1096 1097 1098

[図40]

図40

計算機ノード100A、100B、100Cの決定入力管理テーブル

決定入力総数			1				
送信元情報	シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-A	2A-1	081214	0	0	0	AP1	P1

[図41]

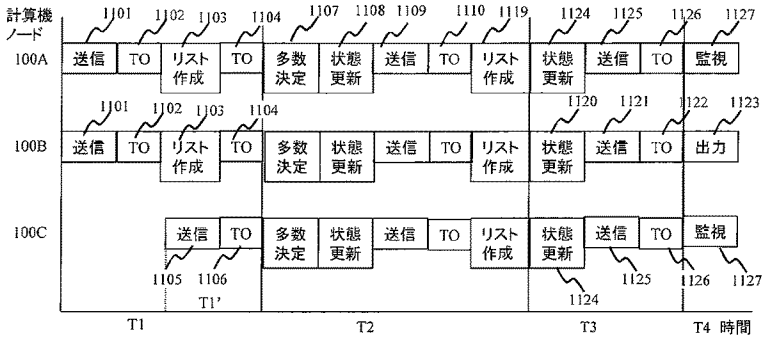
図41

計算機ノード100Bの決定済み入力管理テーブル

決定済み入力総数			1				
送信元情報	シーケンス番号	送信タイムスタップ	順列処理要否フラグ	順列処理ブロック数	順列処理開始番号	コンテンツ	
						アプリ識別子	パラメータ
2-A	2A-1	081214	0	0	0	AP1	P1

[図42]

図42



[図43]

図43

時間別送信に対する各計算機ノードのノード間通信パケット内容

時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				処理状態個別詳細情報
							実行結果
T1	100A	C011	OUTPUT	OUTPUT	0	BEGIN	x
	100B	C011	OUTPUT	OUTPUT	0	BEGIN	x
T1'							
	100C	C011	OUTPUT	OUTPUT	0	BEGIN	x

[図44]

図44

時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				処理状態個別詳細情報		
							ノード識別子	順位	結果
T2	100A	C011	OUTPUT	LEADER	0	BEGIN	100A	2	x
							100B	1	x
							100C	–	-
	100B	C011	OUTPUT	LEADER	0	BEGIN	100A	2	x
							100B	1	x
							100C	–	-
	100C	C011	OUTPUT	LEADER	1	BEGIN	100A	2	x
							100B	1	x
							100C	3	x

[図45]

図45

時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				処理状態個別詳細情報
1151							
1150 T3	100A	C012	OUTPUT	VOTER	0	MONITOR	−
	100B	C012	OUTPUT	VOTER	0	LEADER	−
1152	100C	C012	OUTPUT	VOTER	0	MONITOR	−
1153		1154				1155	

[図46]

図46

時間T1リスト作成に対する計算機ノード100A、100Bのノード間通信パケット内容

時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				処理状態個別詳細情報
1161							実行結果
1160 T1	100A	C011	OUTPUT	OUTPUT	0	BEGIN	X
	100B	C011	OUTPUT	OUTPUT	0	BEGIN	X
1162							
1163							

[図47]

図47

時間T2リスト作成に対する計算機ノード100A、100B、100Cのノード間通信パケット内容

時間	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				処理状態個別詳細情報		
1171							ノード識別子	順位	結果
1170 T2	100A	C011	OUTPUT	LEADER	0	BEGIN	100A	2	x
							100B	1	x
							100C	3	x
1172	100B	C011	OUTPUT	LEADER	0	BEGIN	100A	2	x
							100B	1	x
							100C	3	x
1173	100C	C011	OUTPUT	LEADER	1	BEGIN	100A	2	x
							100B	1	x
							100C	3	x

1174

1175

[図48]

図48

時間別計算機ノード構成管理テーブル変化

317A

時間	入力同期時計算機ノード総数		3			
	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態			
			処理状態識別子	付加情報	Penalty	処理状態
1180 T1	100A	C011	OUTPUT	OUTPUT	0	E
	100B	C011	OUTPUT	OUTPUT	0	E
	100C	C011	OUTPUT	OUTPUT	0	E
1181 T2	100A	C011	OUTPUT	LEADER	0	E
	100B	C011	OUTPUT	LEADER	0	E
	100C	C011	OUTPUT	LEADER	1	E
1182 T3	100A	C012	OUTPUT	VOTER	0	MONITOR
	100B	C012	OUTPUT	VOTER	0	LEADER
	100C	C012	OUTPUT	VOTER	1	MONITOR

[図49]

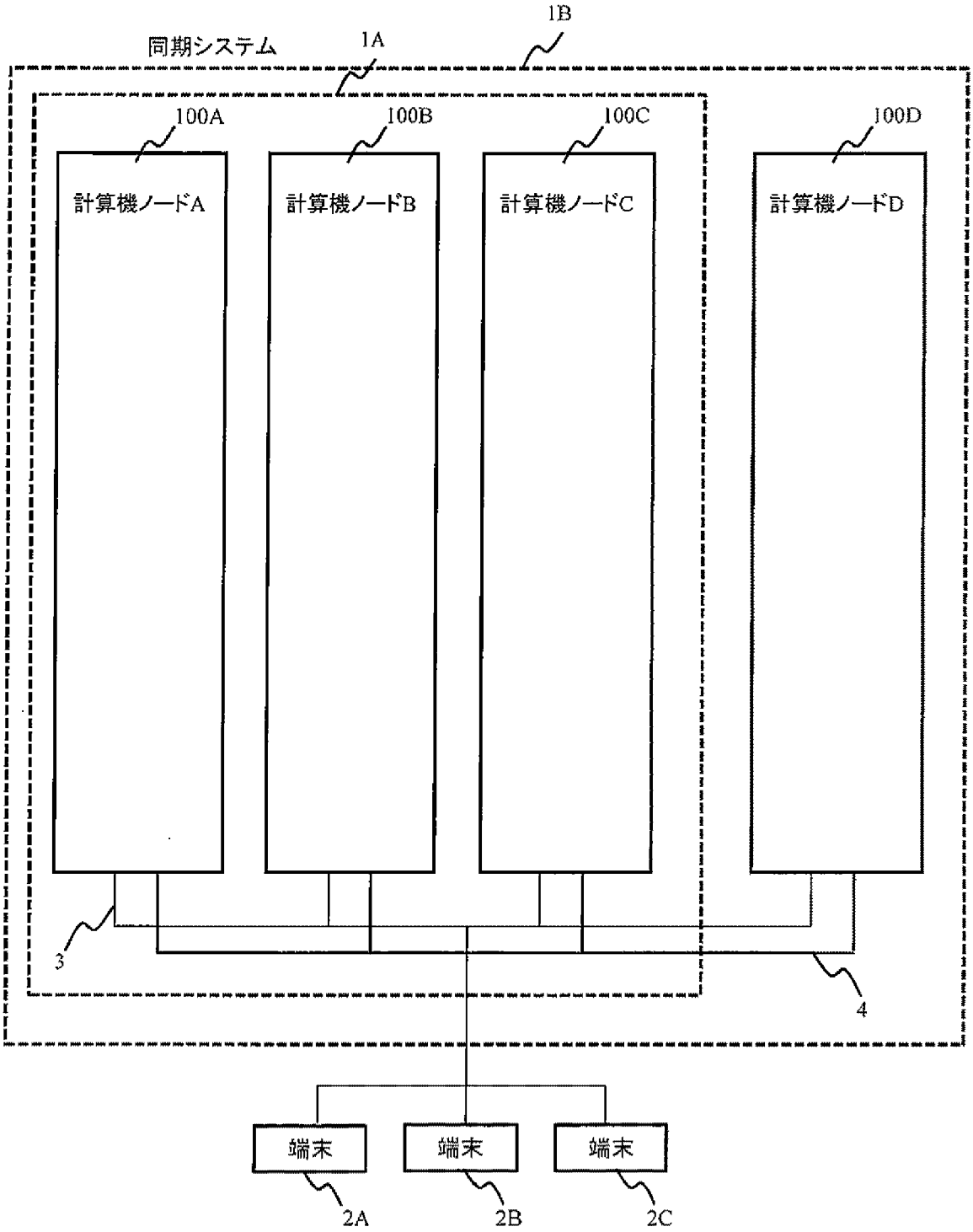
図49

計算機ノード100A、100B、100Cの決定出力管理テーブル

出力同期時計算機ノード総数			3				
順位	計算機ノード識別子	入出力決定処理カウンタ	計算機ノード処理状態				実行結果
			処理状態識別子	付加情報	Penalty	処理状態	
1	100B	C012	OUTPUT	VOTER	0	MONITOR	X
2	100A	C012	OUTPUT	VOTER	0	LEADER	X
3	100C	C012	OUTPUT	VOTER	1	MONITOR	X

[図50]

図50


同期システム
1A
1B
100A
100B
100C
100D
計算機ノードA
計算機ノードB
計算機ノードC
計算機ノードD
3
4
端末
端末
端末
2A
2B
2C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2010/006290

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/18(2006.01)i, *G06F9/52*(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F11/16-11/20, G06F9/46-9/54

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2010
Kokai Jitsuyo Shinan Koho	1971-2010	Toroku Jitsuyo Shinan Koho	1994-2010

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2009-217503 A (Hitachi, Ltd.), 24 September 2009 (24.09.2009), paragraphs [0017] to [0021], [0024] to [0026]; fig. 1, 2B (Family: none)	1-12
A	JP 2003-67214 A (Nippon Telegraph And Telephone Corp.), 07 March 2003 (07.03.2003), paragraphs [0017] to [0032]; fig. 1 to 9 (Family: none)	1-12

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 November, 2010 (18.11.10)	30 November, 2010 (30.11.10)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2010/006290

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2010-113495 A (Nomura Research Institute, Ltd.), 20 May 2010 (20.05.2010), paragraphs [0018] to [0039], [0060] to [0105]; fig. 1, 6 to 15 (Family: none)	1-12
A	JP 2010-198442 A (Toshiba Corp.), 09 September 2010 (09.09.2010), paragraphs [0017] to [0022], [0131]; fig. 1 (Family: none)	1-12

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１０／００６２９０

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F11/18(2006.01)i, G06F9/52(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F11/16－11/20, G06F9/46－9/54

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１０年
日本国実用新案登録公報	１９９６－２０１０年
日本国登録実用新案公報	１９９４－２０１０年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2009-217503 A（株式会社日立製作所）2009.09.24, 段落[0017]－[0021], [0024]－[0026], 第1,2B図 （ファミリーなし）	1－12
A	JP 2003-67214 A（日本電信電話株式会社）2003.03.07, 段落[0017]－[0032], 第1－9図 （ファミリーなし）	1－12

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日 １８．１１．２０１０	国際調査報告の発送日 ３０．１１．２０１０
国際調査機関の名称及びあて先 日本国特許庁（ＩＳＡ／ＪＰ） 郵便番号１００－８９１５ 東京都千代田区霞が関三丁目４番３号	特許庁審査官（権限のある職員）　５Ｂ　３８５６ 北元　健太 電話番号　０３－３５８１－１１０１　内線　３５４５

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Ａ	JP 2010-113495 A (株式会社野村総合研究所) 2010.05.20, 段落[0018]－[0039], [0060]－[0105], 第1,6－15図 （ファミリーなし）	1－12
Ａ	JP 2010-198442 A (株式会社東芝) 2010.09.09, 段落[0017]－[0022], [0131], 第1図 （ファミリーなし）	1－12