

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2006年6月1日 (01.06.2006)

PCT

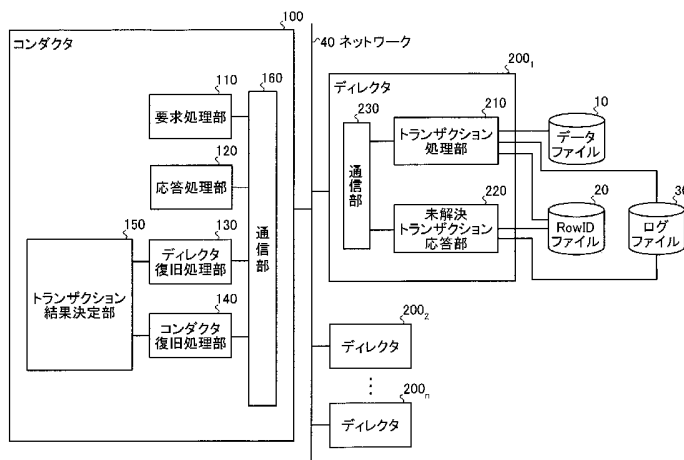
(10) 国際公開番号
WO 2006/057061 A1

- (51) 国際特許分類:
G06F 12/00 (2006.01)
- (21) 国際出願番号: PCT/JP2004/017733
- (22) 国際出願日: 2004年11月29日 (29.11.2004)
- (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): 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 綱川 貴之 (TSUNAKAWA, Takayuki) [JP/JP]; 〒2118588 神奈川
- 県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP). 伊藤 尚洋 (ITO, Naohiro) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).
- (74) 代理人: 酒井 宏明 (SAKAI, Hiroaki); 〒1006019 東京都千代田区霞が関三丁目2番5号 霞が関ビルディング 酒井国際特許事務所 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, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD,

[続葉有]

(54) Title: DISTRIBUTED TRANSACTION PROCESSING METHOD, DEVICE, AND PROGRAM

(54) 発明の名称: 分散トランザクション処理方法、装置、及びプログラム



- 100 CONDUCTOR
- 150 TRANSACTION RESULT DECISION UNIT
- 110 REQUEST PROCESSING UNIT
- 120 RESPONSE PROCESSING UNIT
- 130 DIRECTOR RECOVERY PROCESSING UNIT
- 140 CONDUCTOR RECOVERY PROCESSING UNIT
- 160 COMMUNICATION UNIT
- 40 NETWORK
- 200₁ DIRECTOR
- 230 COMMUNICATION UNIT
- 210 TRANSACTION PROCESSING UNIT
- 220 UNSOLVED TRANSACTION RESPONSE UNIT
- 10 DATA FILE
- 20 RowID FILE
- 30 LOG FILE
- 200₂ DIRECTOR
- 200_n DIRECTOR

(57) Abstract: For a commit request from an application, a request processing unit transmits a commit processing request to a director concerned. When a response processing unit has received a response from any one of the directors, the successful commit is reported to the application. When the director is downed and a transaction not solved is present, upon recovery of the director, a director recovery processing unit instructs the recovered director to process the unsolved transaction according to the processing result of the unsolved transaction by other director. Moreover, when the conductor is downed while transmitting a commit processing request to the director and the commit processing request could not be transmitted to some directors, after recovery by the conductor recovery processing unit, an inquiry about an unsolved transaction is performed to each director. The director which has responded is instructed to execute processing of the unsolved transaction according to the processing result of other director for the unsolved transaction responded by the director.

(57) 要約: アプリケーションからのコミット要求に対して関連するディレクタに要求処理部がコミット処理要求を送信し、応答処理部がいずれかのディレクタから応答を受け取った場合にアプリケーションにコミットの成功を通知し、ディレクタがダウンして未解決トランザクションがある場合には、ディレクタが復旧した際に、ディレ

クタ復旧処理部が他のディレクタによる未解決トランザクションの処理結果に基づいて未解決トランザクションの処理を

[続葉有]

WO 2006/057061 A1



SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, YU, ZA, ZM, ZW.

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

(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, 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 (AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,
IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI

添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される
各PCTガゼットの巻頭に掲載されている「コードと略語
のガイダンスノート」を参照。

復旧したディレクタに指示する。また、ディレクタにコミット処理要求を送信中にコンダクタがダウンして一部の
ディレクタにコミット処理要求が送信できなかった場合には、コンダクタ復旧処理部が復旧後に各ディレクタに対
して未解決トランザクションの問合せを行い、ディレクタが応答した未解決トランザクションに対する他のディレ
クタの処理結果に基づいて未解決トランザクションの処理を、応答したディレクタに指示する。

明 細 書

分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置

技術分野

[0001] 本発明は、装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置に関し、特に、コミット処理を高速に行うとともに、導入・運用の負担を軽減することができる分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置に関するものである。

背景技術

[0002] 従来、一つのトランザクションを処理する場合に複数のデータベースに対してアクセスする必要がある分散トランザクション処理では、複数のデータベース間の整合性を保証するために、2フェーズコミットが用いられてきた。

[0003] 2フェーズコミットでは、コミット処理を2段階のフェーズにわけ、第1段階でデータベースに対してコミットの準備ができていないか否かを確認し(プリペア)、全てのデータベースからコミットの準備ができたことが確認できた場合にのみ第2段階でコミットを実行し、それ以外の場合には第2段階でロールバックを行うことによってデータベース間の整合性を保証している(例えば、非特許文献1参照。)

[0004] 非特許文献1:「分散トランザクションに挑戦しよう!」、[平成16年10月12日検索]、インターネット<URL: HYPERLINK
"http://www.microsoft.com/japan/enable/training/kblight/t004/7/01.htm"
http://www.ogis-ri.co.jp/otc/hiroba/technical/DTP/step2>

発明の開示

発明が解決しようとする課題

[0005] しかしながら、2フェーズコミットでは、コミット処理を2段階のフェーズにわけて行うた

めに、処理効率が悪いという問題がある。また、分散トランザクション処理を管理するトランザクションマネージャを実行するコンピュータに障害が発生した場合にトランザクションマネージャが復旧処理で使用するトランザクションログが必要であり、トランザクションログのための導入・運用の負担が大きいという問題もある。

[0006] 本発明は、上記に鑑みてなされたものであって、コミット処理を高速に行うとともに、導入・運用の負担を軽減することができる分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置を提供することを目的とする。

課題を解決するための手段

[0007] 上述した課題を解決し、目的を達成するため、請求項1の発明に係る分散トランザクション処理プログラムは、装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理プログラムであって、前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手順と、前記コミット要求手順により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手順と、前記コミット要求手順により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示手順と、をコンピュータに実行させることを特徴とする。

[0008] この請求項1の発明によれば、アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求し、要求したコミット処理に対していずれかの処理装置から応答があった場合にアプリケーションに対してコミットの成功を応答し、要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて故障処理装置に対して指示するよう構成したので、1フェーズコミットによってデータベース

間の整合性を維持することができる。

- [0009] また、請求項2の発明に係る分散トランザクション処理プログラムは、請求項1の発明において、前記復旧指示手順は、前記故障処理装置が復旧したか否かを判定する復旧判定手順と、前記復旧判定手順により故障処理装置が復旧したと判定された場合に、故障処理装置の未解決トランザクションの処理結果について他の処理装置に問合せを行い、問合せ結果に基づいて未解決トランザクションに対する処理を故障処理装置に要求する処理装置復旧処理手順とを、コンピュータに実行させることを特徴とする。
- [0010] この請求項2の発明によれば、故障処理装置が復旧したか否かを判定し、故障処理装置が復旧したと判定した場合に、故障処理装置の未解決トランザクションの処理結果について他の処理装置に問合せを行い、問合せ結果に基づいて未解決トランザクションに対する処理を故障処理装置に要求するよう構成したので、処理装置によるトランザクションログの管理だけが必要であり、当該分散トランザクション処理プログラムによるトランザクションログの管理を不要とすることができる。
- [0011] また、請求項3の発明に係る分散トランザクション処理プログラムは、請求項2の発明において、前記復旧判定手順は、前記故障処理装置に対して定期的に接続を確認することによって該故障処理装置が復旧したか否かを判定することを特徴とする。
- [0012] この請求項3の発明によれば、故障処理装置に対して定期的に接続を確認することによって故障処理装置が復旧したか否かを判定するよう構成したので、確実に故障処理装置の復旧を検知することができる。
- [0013] また、請求項4の発明に係る分散トランザクション処理プログラムは、請求項1、2または3の発明において、自装置の故障からの復旧時に処理装置に対して未解決トランザクションを問い合わせる未解決問合せを行い、該未解決問合せに対して処理装置から応答された未解決トランザクションの処理結果について他の処理装置に問い合わせる処理結果問合せを行う未解決トランザクション処理結果問合せ手順と、前記未解決トランザクション処理結果問合せ手順による処理結果問合せに対する応答結果に基づいて該未解決トランザクションに対する処理を未解決問合せに対して未解決トランザクションを応答した処理装置に要求する自装置復旧処理手順とを、さらに

コンピュータに実行させることを特徴とする。

- [0014] この請求項4の発明によれば、自装置の故障からの復旧時に処理装置に対して未解決トランザクションを問い合わせる未解決問合せを行い、未解決問合せに対して処理装置から応答された未解決トランザクションの処理結果について他の処理装置に問い合わせる処理結果問合せを行い、処理結果問合せに対する応答結果に基づいて未解決トランザクションに対する処理を未解決問合せに対して未解決トランザクションを応答した処理装置に要求するよう構成したので、当該分散トランザクション処理プログラムを実行するコンピュータが一部の処理装置に対してコミット処理を要求した後に故障した場合にも、復旧後に残りの処理装置のコミット処理を正しく行うことができる。
- [0015] また、請求項5の発明に係る分散トランザクション処理プログラムは、請求項1の発明において、前記コミット要求手順は、コミット処理の要求がいずれの処理装置にも受け付けられなかった場合に、前記アプリケーションに対してコミットの失敗を応答することを特徴とする。
- [0016] この請求項5の発明によれば、コミット処理の要求がいずれの処理装置にも受け付けられなかった場合に、アプリケーションに対してコミットの失敗を応答するよう構成したので、コミット処理を要求した全ての処理装置が故障している場合にもアプリケーションに迅速に応答することができる。
- [0017] また、請求項6の発明に係る分散トランザクション処理プログラムは、請求項1の発明において、前記コミット応答手順は、前記コミット要求手順により要求したコミット処理に対していずれの処理装置からも応答がなかった場合には、前記アプリケーションにコミット結果を不明と応答することを特徴とする。
- [0018] この請求項6の発明によれば、要求したコミット処理に対していずれの処理装置からも応答がなかった場合には、アプリケーションにコミット結果を不明と応答するよう構成したので、コミット処理を要求した全ての処理装置が要求受付後に故障した場合にもアプリケーションに確実に応答することができる。
- [0019] また、請求項7の発明に係る分散トランザクション処理方法は、装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置

にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理方法であって、前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求工程と、前記コミット要求工程により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答工程と、前記コミット要求工程により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示工程と、を含んだことを特徴とする。

[0020] この請求項7の発明によれば、アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求し、要求したコミット処理に対していずれかの処理装置から応答があった場合にアプリケーションに対してコミットの成功を応答し、要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて故障処理装置に対して指示するよう構成したので、1フェーズコミットによってデータベース間の整合性を維持することができる。

[0021] また、請求項8の発明に係る分散トランザクション処理装置は、装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理装置であって、前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手段と、前記コミット要求手段により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手段と、前記コミット要求手段により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結

果に基づいて該故障処理装置に対して指示する復旧指示手段と、を備えたことを特徴とする。

[0022] この請求項8の発明によれば、アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求し、要求したコミット処理に対していずれかの処理装置から応答があった場合にアプリケーションに対してコミットの成功を応答し、要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて故障処理装置に対して指示するよう構成したので、1フェーズコミットによってデータベース間の整合性を維持することができる。

[0023] また、請求項9の発明に係る分散トランザクション処理装置は、一つのトランザクションを分散処理する複数の処理装置と、アプリケーションからトランザクション処理要求を受け付けて該複数の処理装置に分散処理を要求するトランザクション管理装置とから構成される分散トランザクション処理装置であって、前記トランザクション管理装置は、前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手段と、前記コミット要求手段により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手段と、前記コミット要求手段により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示手段と、を備え、前記処理装置は、装置自身の故障以外では前記トランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行って該トランザクション管理装置に応答することができるコミット処理手段と、故障から復旧した際に、前記トランザクション管理装置からの指示に基づいて前記未解決トランザクションの処理を行う復旧処理手段と、を備えたことを特徴とする。

[0024] この請求項9の発明によれば、トランザクション管理装置は、アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求し、要求したコミット処理に

対していずれかの処理装置から応答があった場合にアプリケーションに対してコミットの成功を応答し、要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて故障処理装置に対して指示し、処理装置は、装置自身の故障以外ではトランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行ってトランザクション管理装置に応答することができ、故障から復旧した際に、トランザクション管理装置からの指示に基づいて未解決トランザクションの処理を行うよう構成したので、1フェーズコミットによってデータベース間の整合性を維持することができる。

- [0025] また、請求項10の発明に係る分散トランザクション処理方法は、一つのトランザクションを分散処理する複数の処理装置と、アプリケーションからトランザクション処理要求を受け付けて該複数の処理装置に分散処理を要求するトランザクション管理装置とから構成される分散トランザクション処理装置のコミット処理方法であって、前記トランザクション管理装置が前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求工程と、前記処理装置が装置自身の故障以外では前記トランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行って該トランザクション管理装置に応答するコミット処理工程と、前記コミット要求工程により要求したコミット処理に対して、前記トランザクション管理装置が、いずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答工程と、前記コミット要求工程により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、前記トランザクション管理装置が、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示工程と、故障から復旧した際に、前記故障処理装置が、前記トランザクション管理装置からの指示に基づいて前記未解決トランザクションの処理を行う復旧処理工程と、を含んだことを特徴とする。

[0026] この請求項10の発明によれば、トランザクション管理装置がアプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求し、処理装置が装置自身の故障以外ではトランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行ってトランザクション管理装置に応答し、要求したコミット処理に対して、トランザクション管理装置が、いずれかの処理装置から応答があった場合にアプリケーションに対してコミットの成功を応答し、要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、トランザクション管理装置が、応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて故障処理装置に対して指示し、故障から復旧した際に、故障処理装置が、トランザクション管理装置からの指示に基づいて未解決トランザクションの処理を行うよう構成したので、1フェーズコミットによってデータベース間の整合性を維持することができる。

発明の効果

[0027] 請求項1、7、8、9および10の発明によれば、1フェーズコミットによってデータベース間の整合性を維持することができるので、コミット処理を高速に行うことができるという効果を奏する。

[0028] また、請求項2の発明によれば、処理装置によるトランザクションログの管理だけが必要であり、当該分散トランザクション処理プログラムによるトランザクションログの管理を不要とするので、導入・運用の負担を軽減することができるという効果を奏する。

[0029] また、請求項3の発明によれば、確実に故障処理装置の復旧を検知するので、データベース間の整合性を確実に維持することができるという効果を奏する。

[0030] また、請求項4の発明によれば、当該分散トランザクション処理プログラムを実行するコンピュータが一部の処理装置に対してコミット処理を要求した後に故障した場合にも、復旧後に残りの処理装置のコミット処理を正しく行うことができるので、1フェーズコミットによってデータベース間の整合性を維持することができるという効果を奏する。

[0031] また、請求項5の発明によれば、コミット処理を要求した全ての処理装置が故障している場合にアプリケーションに迅速に応答するので、コミット処理を高速に行うことが

できるという効果を奏する。

- [0032] また、請求項6の発明によれば、コミット処理を要求した全ての処理装置が要求受付後に故障した場合にもアプリケーションに確実に応答することができるので、アプリケーションは適切な復旧動作行うことができるという効果を奏する。

図面の簡単な説明

- [0033] [図1]図1は、本実施例に係る分散トランザクション処理システムのシステム構成を示す機能ブロック図である。
- [図2]図2は、各レコードに対応してRowIDファイルに作成されるエントリのデータ構造の一例を示す図である。
- [図3]図3は、ログファイルに含まれる未解決トランザクション一覧のデータ構造の一例を示す図である。
- [図4]図4は、コンダクタによるコミット処理の処理手順を示すフローチャートである。
- [図5]図5は、図1に示したディレクタ復旧処理部によるディレクタ復旧処理の処理手順を示すフローチャートである。
- [図6]図6は、図1に示したコンダクタ復旧処理部によるコミット復旧処理の処理手順を示すフローチャートである。
- [図7]図7は、図1に示したトランザクション結果決定部によるトランザクション結果決定処理の処理手順を示すフローチャートである。
- [図8]図8は、2台のディレクタが分散トランザクションを処理する場合に、ディレクタがダウンする各タイミングに対応するデータベース整合性維持動作を示す図である。
- [図9]図9は、本実施例に係るコンダクタプログラムおよびディレクタプログラムを実行するコンピュータを示す図である。

符号の説明

- [0034] 10 データファイル
20 RowIDファイル
30 ログファイル
40 ネットワーク
100 コンダクタ

- 110 要求処理部
- 120 応答処理部
- 130 ディレクタ復旧処理部
- 140 コンダクタ復旧処理部
- 150 トランザクション結果決定部
- 160 通信部
- 200₁ ~ 200_n ディレクタ
- 210 トランザクション処理部
- 220 未解決トランザクション応答部
- 230 通信部
- 300 コンピュータ
- 310 HDD
- 311 コンダクタプログラム
- 312 ディレクタプログラム
- 320 RAM
- 330 ROM
- 340 CPU
- 341 コンダクタプロセス
- 342 ディレクタプロセス
- 350 LANインタフェース
- 360 I/Oインタフェース

発明を実施するための最良の形態

[0035] 以下に、本発明に係る分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置の実施例を図面に基づいて詳細に説明する。なお、この実施例によりこの発明が限定されるものではない。

実施例

[0036] まず、本実施例に係る分散トランザクション処理システムの構成について説明する。図1は、本実施例に係る分散トランザクション処理システムのシステム構成を示す機

能ブロック図である。

- [0037] 同図に示すように、この分散トランザクション処理システムは、アプリケーションからトランザクションを受け付けるコンダクタ100と、コンダクタ100によって受け付けられたトランザクションを分散処理するn台のディレクタ 200_1 〜 200_n がネットワーク40を介して接続されて構成される。
- [0038] コンダクタ100は、ネットワーク40を介してアプリケーションから受け付けたトランザクションの分散処理を関係するディレクタに指示し、各ディレクタの処理結果に基づいてトランザクションの処理結果をアプリケーションに応答する装置である。
- [0039] なお、この分散トランザクション処理システムでは、コンダクタ100は、1フェーズコミットによってコミット処理を行う。また、ディレクタ 200_1 〜 200_n は、故障によってコミット処理を行えない場合以外は、コミット処理要求を受け付けると必ずコミット処理を行うことができるものとする。
- [0040] コンダクタ100は、要求処理部110と、応答処理部120と、ディレクタ復旧処理部130と、コンダクタ復旧処理部140と、トランザクション結果決定部150と、通信部160とを有する。
- [0041] 要求処理部110は、アプリケーションからトランザクションを受け付け、受け付けたトランザクションの分散処理を関係するディレクタに指示する処理部である。この要求処理部110は、アプリケーションからコミット要求を受け付けた場合には、関係するディレクタにコミット処理を要求する。
- [0042] ここで、関係する全てのディレクタがダウンしていて、いずれのディレクタにもコミット処理を依頼することができない場合には、この要求処理部110は、アプリケーションにコミットの失敗を通知する。
- [0043] 応答処理部120は、要求処理部110によりトランザクションの分散処理を指示されたディレクタから分散トランザクション処理結果を受け取り、各ディレクタの分散トランザクション処理結果に基づいてアプリケーションにトランザクション処理結果を通知する処理部である。
- [0044] この応答処理部120は、要求処理部110によりコミット処理を指示されたディレクタのうちのいずれかのディレクタから応答を受け取ると、コミットが成功したことをアプリケ

ーションに通知する。

- [0045] また、この応答処理部120は、要求処理部110によりコミット処理を指示されたいずれのディレクタからも所定の時間内に応答がない場合、すなわち、コミット処理を指示された全てのディレクタがコミット処理の指示を受信後にダウンした場合には、アプリケーションにコミット処理の結果が不明であることを通知する。
- [0046] ディレクタ復旧処理部130は、ディレクタ $200_1 \sim 200_n$ の中にダウンしたディレクタがある場合に、ダウンしたディレクタが復旧したか否かを監視し、ディレクタが復旧した際に復旧処理を行う処理部である。
- [0047] 具体的には、このディレクタ復旧処理部130は、復旧したディレクタに対してコミット処理が行われていない未解決トランザクションの問合せを行い、未解決トランザクションがある場合には、トランザクション結果決定部150を用いてその未解決トランザクションに対して行うべき処理を決定し、決定した処理を復旧したディレクタに指示する。
- [0048] このディレクタ復旧処理部130が、復旧したディレクタに対してコミット処理が行われていない未解決トランザクションの問合せを行い、未解決トランザクションがある場合には、トランザクション結果決定部150を用いてその未解決トランザクションに対して行うべき処理を決定し、決定した処理を復旧したディレクタに指示することによって、コミット処理の際にダウンしたディレクタがある場合にもデータベース間の整合性を保証することができる。
- [0049] すなわち、応答処理部120がいずれかのディレクタから応答があった場合に直ちにアプリケーションに対してコミットの成功を通知し、残りのディレクタにコミット処理要求受信後コンダクタ100への応答前にダウンしたディレクタがある場合にも、そのディレクタが復旧後コミット処理を行うことができ、データベース間の整合性を保証することができる。
- [0050] コンダクタ復旧処理部140は、コンダクタ100がダウンして復旧した際に、復旧処理を行う処理部である。このコンダクタ復旧処理部140は、復旧処理の一部として、未解決トランザクションの処理を行う。
- [0051] 具体的には、このコンダクタ復旧処理部140は、全てのディレクタ $200_1 \sim 200_n$ に対して未解決トランザクションの問合せを行い、あるディレクタが未解決トランザクション

があることを応答した場合には、トランザクション結果決定部150を用いてその未解決トランザクションに対して行うべき処理を決定し、決定した処理を未解決トランザクションを応答したディレクタに指示する。

[0052] このコンダクタ復旧処理部140が、復旧処理の一部として、全てのディレクタ200₁〜200_nに対して未解決トランザクションの問合せを行い、あるディレクタが未解決トランザクションがあることを応答した場合には、トランザクション結果決定部150を用いてその未解決トランザクションに対して行うべき処理を決定し、決定した処理を未解決トランザクションを応答したディレクタに指示することによって、ディレクタ200₁〜200_nの一部にコミット要求を送信してコンダクタ100がダウンした場合にも、データベース間の整合性を保証することができる。

[0053] トランザクション結果決定部150は、ディレクタ復旧処理部130およびコンダクタ復旧処理部140の指示にしたがって、未解決トランザクションに対して行うべき処理を決定する処理部である。

[0054] 具体的には、このトランザクション結果決定部150は、未解決トランザクションに対するコミット結果を他のディレクタに問い合わせ、他のディレクタのコミット結果に基づいて未解決トランザクションに対して行うべき処理を決定する。

[0055] すなわち、このトランザクション結果決定部150は、未解決トランザクションに対して他のディレクタがコミット処理を行った場合には、未解決トランザクションに対して行うべき処理をコミット処理に決定し、未解決トランザクションに対して他のディレクタがロールバックを行った場合には、未解決トランザクションに対して行うべき処理をロールバックに決定する。

[0056] 通信部160は、アプリケーションを実行するコンピュータやディレクタ200₁〜200_nなどとネットワーク40を介して通信する処理部であり、アプリケーションからのコミット要求やディレクタ200₁〜200_nからの応答などを受信し、アプリケーションへのコミット結果の応答やディレクタ200₁〜200_nへのコミット処理要求などを送信する。

[0057] ディレクタ200₁〜200_nは、分散トランザクションを処理するためにアクセスする必要がある複数のデータベースのそれぞれを管理する装置であり、コンダクタ100からトランザクション処理要求を受け付けてトランザクションを分散処理する。なお、ディレクタ

200₁ ~ 200_n はいずれも同様の構成を有するので、ここではディレクタ200₁ を例にとって説明する。

- [0058] ディレクタ200₁ は、トランザクション処理部210と、未解決トランザクション応答部220と、通信部230とを有する。
- [0059] トランザクション処理部210は、コンダクタ100からトランザクション処理要求を受け付けてトランザクションを分散処理し、レコード形式でデータを格納するデータファイル10を更新する処理部である。
- [0060] また、このトランザクション処理部210は、データファイル10のレコードの格納位置を記憶するRowIDファイル20およびトランザクションログを記憶するログファイル30を更新する。
- [0061] 図2は、各レコードに対応してRowIDファイル20に作成されるエントリのデータ構造の一例を示す図である。同図に示すように、このエントリには、レコード格納位置とトランザクションIDとが含まれる。
- [0062] レコード格納位置は、データファイル10のレコードの格納位置を示す。トランザクションIDは、このエントリに対応するレコードの追加・削除・更新を最後に行ったトランザクションを識別する識別子であり、トランザクションの処理が完了するとトランザクション処理部210によって更新される。
- [0063] 図3は、ログファイル30に含まれる未解決トランザクション一覧のデータ構造の一例を示す図である。同図に示すように、この未解決トランザクション一覧には、コミット処理が行われていないm個のトランザクションの識別子であるトランザクションID₁ ~ トランザクションID_m が記憶されている。
- [0064] 未解決トランザクション応答部220は、コンダクタ100からの未解決トランザクションの問合せに対してログファイル30の未解決トランザクション一覧に記憶されたトランザクションIDをコンダクタ100に応答する処理部である。
- [0065] この未解決トランザクション応答部220が、コンダクタ100からの問合せに対して未解決トランザクションを応答することによって、コンダクタ100は、このディレクタ200₁ またはコンダクタ100自身の復旧処理において処理が必要な未解決トランザクションを特定することができる。

- [0066] また、この未解決トランザクション応答部220は、コンダクタ100からの他のディレクタの未解決トランザクションに関するコミット結果の問合せに対して、ログファイル30の未解決トランザクション一覧およびRowIDファイル20を用いてコンダクタ100に応答する。
- [0067] 具体的には、この未解決トランザクション応答部220は、コンダクタ100から未解決トランザクションのトランザクションIDを受け取り、そのトランザクションIDが未解決トランザクション一覧にあればコミット未完了であることを応答し、未解決トランザクション一覧になければRowIDファイル20内にあるか否かを調べ、RowIDファイル20内にある場合には、トランザクションがコミットされているのでコミット完了を応答し、RowIDファイル20内にはない場合には、トランザクションがコミットされていないのでコミット未完了を応答する。
- [0068] この未解決トランザクション応答部220が、コンダクタ100からの他のディレクタの未解決トランザクションに関するコミット結果の問合せに対して、ログファイル30の未解決トランザクション一覧およびRowIDファイル20を用いてコミットが完了したか否かをコンダクタ100に応答することによって、コンダクタ100は、他のディレクタの未解決トランザクションに対する処理を決定することができる。
- [0069] 通信部230は、コンダクタ100などとネットワーク40を介して通信する処理部であり、例えば、コンダクタ100からコミット処理要求や未解決トランザクションの問合せを受信し、コンダクタ100へコミット結果や未解決トランザクションのトランザクションIDを送信する。
- [0070] 次に、コンダクタ100によるコミット処理の処理手順について説明する。図4は、コンダクタ100によるコミット処理の処理手順を示すフローチャートである。同図に示すように、このコンダクタ100は、要求処理部110がアプリケーションからコミット要求を受信し(ステップS101)、関連するディレクタ群にコミット処理要求を送信する(ステップS102)。
- [0071] そして、いずれかのディレクタにコミット処理要求を送信できたか否かを判定し(ステップS103)、いずれのディレクタにもコミット処理要求を送信できなかった場合には、全ディレクタがダウンしている場合であるので、アプリケーションにコミットの失敗を通

知する(ステップS108)。

[0072] 一方、いずれかのディレクタにコミット処理要求を送信できた場合には、応答処理部120が、いずれかのディレクタから応答があると(ステップS104、肯定)、アプリケーションへコミットの成功を通知し(ステップS105)、いずれのディレクタからも応答がない場合には、タイムアウトであるか否かを判定する(ステップS106)。

[0073] その結果、タイムアウトでない場合には、引き続きディレクタからの応答を待ち、タイムアウトの場合には、コミット処理要求を受け付けた全てのディレクタが要求受付後にダウンした場合であり、各ディレクタでコミット処理がどこまで行われたかわからないので、アプリケーションへコミット結果が不明であることを通知する(ステップS107)。

[0074] このように、コンダクタ100がディレクタにコミット処理要求を送信し、いずれかのディレクタから応答があると直ちにアプリケーションにコミットの成功を通知するので、アプリケーションからのコミット要求を高速に処理することができる。

[0075] 次に、図1に示したディレクタ復旧処理部130によるディレクタ復旧処理の処理手順について説明する。図5は、図1に示したディレクタ復旧処理部130によるディレクタ復旧処理の処理手順を示すフローチャートである。

[0076] 同図に示すように、このディレクタ復旧処理部130は、ダウンしたディレクタがある場合に、ダウンしたディレクタに定期的に接続を試行し(ステップS201)、接続できると、未解決トランザクションの問合せを行う(ステップS202)。

[0077] そして、未解決トランザクションがあるか否かを判定し(ステップS203)、未解決トランザクションがある場合には、未解決トランザクションを一つ選択し(ステップS204)、選択した未解決トランザクションに対して行うべき処理を決定するトランザクション結果決定処理を行う(ステップS205)。

[0078] そして、未解決トランザクションに対して行うべき処理が決定したか否か、すなわち、コミットまたはロールバックが決定したか否かを判定し(ステップS206)、コミットかロールバックが決定した場合には、復旧したディレクタに対して決定結果に基づいて未解決トランザクションを処理するように指示し(ステップS207)、ステップS203に戻って次の未解決トランザクションの処理に移る。

[0079] 一方、コミットかロールバックかが不明な場合には、復旧したディレクタにその未解

決トランザクションについては指示することなく、ステップS203に戻って次の未解決トランザクションの処理に移る。

- [0080] このように、このディレクタ復旧処理部130が、復旧したディレクタの未解決トランザクションの処理をトランザクション結果決定処理によって決定し、決定した処理を復旧したディレクタに指示することによって、コミット処理前にダウンしたディレクタのコミット処理を復旧後に正しく行うことができる。
- [0081] 次に、図1に示したコンダクタ復旧処理部140によるコミット復旧処理の処理手順について説明する。図6は、図1に示したコンダクタ復旧処理部140によるコミット復旧処理の処理手順を示すフローチャートである。
- [0082] 同図に示すように、このコンダクタ復旧処理部140は、ダウンから復旧後にディレクタを一つ選択し(ステップS301)、未解決トランザクションの問合せを行う(ステップS302)。
- [0083] そして、未解決トランザクションがあるか否かを判定し(ステップS303)、未解決トランザクションがない場合には、全ディレクタを処理したか否かを判定し(ステップS304)、全ディレクタを処理していない場合には、ステップS301に戻って次のディレクタの処理に移り、全ディレクタを処理した場合には、処理を終了する。
- [0084] 一方、未解決トランザクションがある場合には、未解決トランザクションを一つ選択し(ステップS305)、選択した未解決トランザクションに対して行うべき処理を決定するトランザクション結果決定処理を行う(ステップS306)。
- [0085] そして、未解決トランザクションに対して行うべき処理が決定したか否か、すなわち、コミットまたはロールバックが決定したか否かを判定し(ステップS307)、コミットかロールバックが決定した場合には、選択しているディレクタに対して決定結果に基づいて未解決トランザクションを処理するように指示し(ステップS308)、ステップS303に戻って次の未解決トランザクションの処理に移る。
- [0086] 一方、コミットかロールバックかが不明な場合には、選択しているディレクタにその未解決トランザクションについては指示することなく、ステップS303に戻って次の未解決トランザクションの処理に移る。
- [0087] このように、このコンダクタ復旧処理部140がディレクタに未解決トランザクションの

問合せを行い、未解決トランザクションに対して行うべき処理をトランザクション結果決定処理によって決定し、決定した処理をディレクタに指示することによって、コミット要求を一部のディレクタに送信した後にコンダクタ100がダウンした場合にも、復旧後に残りのディレクタのコミット処理を正しく行うことができる。

- [0088] 次に、図1に示したトランザクション結果決定部150によるトランザクション結果決定処理の処理手順について説明する。図7は、図1に示したトランザクション結果決定部150によるトランザクション結果決定処理の処理手順を示すフローチャートである。なお、このトランザクション結果決定処理は、図5のステップS205の処理および図6のステップS306の処理に対応する。
- [0089] 図7に示すように、このトランザクション結果決定部150は、未解決トランザクションの状態が未完了であるか否かを判定し(ステップS401)、未解決トランザクションの状態が未完了である場合には、結果を未確定として処理を終了する(ステップS412)。
- [0090] 一方、未解決トランザクションの状態が未完了でない場合には、ダウンディレクタの数を数えるための変数としてダウンディレクタ数を「0」に初期化し(ステップS402)、ディレクタを一つ選択する(ステップS403)。
- [0091] そして、選択したディレクタに対して未解決トランザクションの処理結果の問合せを行い(ステップS404)、ディレクタがダウンしているか否かを判定する(ステップS405)。
- [0092] その結果、ディレクタがダウンしていない場合には、未解決トランザクションの処理結果の問合せに対する応答がコミットであるか否かを判定し(ステップS406)、コミットである場合には、結果をコミットとして処理を終了する(ステップS407)。
- [0093] 一方、未解決トランザクションの処理結果の問合せに対する応答がコミットでない場合には、未解決トランザクションの処理結果は不明であり、他のディレクタに問い合わせる必要があるので、ステップS409に進んで全ディレクタに対して問い合わせたか否かを判定する。
- [0094] また、ディレクタがダウンしている場合には、ダウンディレクタ数に「1」を加え(ステップS408)、全ディレクタに対して問い合わせたか否かを判定する(ステップS409)。
- [0095] その結果、全ディレクタに対して問い合わせしていない場合には、ステップS403に戻

って次のディレクタを選択し、全ディレクタに対して問い合わせた場合には、ダウンディレクタ数が「0」であるか否かを判定する(ステップS410)。

- [0096] そして、ダウンディレクタ数が「0」である場合には、全てのディレクタがコミットを行っていないので、結果をロールバックとして処理を終了し(ステップS411)、ダウンディレクタ数が「0」でない場合には、ダウンしているディレクタによる未解決トランザクションの処理結果が不明であるので、結果を未確定として処理を終了する(ステップS412)。
- [0097] このように、このトランザクション結果決定部150が、未解決トランザクションのコミット結果をディレクタに順に問い合わせ、コミットしたディレクタがある場合には、未解決トランザクションの結果をコミットと決定し、全ディレクタがコミットしなかった場合には、未解決トランザクションの結果をロールバックと決定することによって、コンダクタ100によるトランザクション処理結果のログ記録をなくすることができる。
- [0098] 次に、2台のディレクタが分散トランザクションを処理する場合に、ディレクタがダウンする各タイミングに対応するデータベース整合性維持動作について説明する。図8は、2台のディレクタが分散トランザクションを処理する場合に、ディレクタがダウンする各タイミングに対応するデータベース整合性維持動作を示す図である。
- [0099] 同図に示すように、2台のディレクタ D_1 および D_2 がいずれもダウンしない場合には、コミット要求に対するアプリケーションへの応答はコミット成功となり、トランザクションの結果はコミットとなる。また、データベース間の整合性を維持するための動作は何も必要ない。
- [0100] また、ディレクタ D_1 がダウンせずディレクタ D_2 がダウンした場合には、ディレクタ D_2 がダウンしたタイミングに関係なく、コミット要求に対するアプリケーションへの応答はコミット成功となり、トランザクションの結果はコミットとなる。
- [0101] ただし、 D_2 がコミット指示を受け付ける前またはコミット処理中にダウンした場合には、 D_2 は再起動後、コンダクタ100を通じて D_1 でのコミット完了を知り、コミットする。また、 D_2 がコミット応答前にダウンした場合には、データベース間の整合性を維持するための動作は何も必要ない。
- [0102] また、ディレクタ D_1 がコミット指示を受け付ける前にダウンし、ディレクタ D_2 もコミット指

示を受け付ける前にダウンした場合には、コミット要求に対するアプリケーションへの応答はコミット失敗となり、トランザクションの結果はロールバックとなり、D₁ および D₂ は再起動後、コンダクタ100を通じてコミットが完了したディレクタがなかったことを知り、ロールバックする。

[0103] また、ディレクタD₁ がコミット指示を受け付ける前にダウンし、ディレクタD₂ がコミット処理中にダウンした場合には、コミット要求に対するアプリケーションへの応答は不明となり、トランザクションの結果はロールバックとなり、D₁ および D₂ は再起動後、コンダクタ100を通じてコミットが完了したディレクタがなかったことを知り、ロールバックする。

[0104] また、ディレクタD₁ がコミット指示を受け付ける前にダウンし、ディレクタD₂ がコミット応答前にダウンした場合には、コミット要求に対するアプリケーションへの応答は不明となり、トランザクションの結果はコミットとなり、D₁ は再起動後、コンダクタ100を通じてD₂ でのコミット完了を知り、コミットする。

[0105] また、ディレクタD₁ がコミット処理中にダウンし、ディレクタD₂ もコミット処理中にダウンした場合には、コミット要求に対するアプリケーションへの応答は不明となり、トランザクションの結果はロールバックとなり、D₁ および D₂ は再起動後、コンダクタ100を通じてコミットが完了したディレクタがなかったことを知り、ロールバックする。

[0106] また、ディレクタD₁ がコミット処理中にダウンし、ディレクタD₂ がコミット応答前にダウンした場合には、コミット要求に対するアプリケーションへの応答は不明となり、トランザクションの結果はコミットとなり、D₁ は再起動後、コンダクタ100を通じてD₂ でのコミット完了を知り、コミットする。

[0107] また、ディレクタD₁ がコミット応答前にダウンし、ディレクタD₂ もコミット応答前にダウンした場合には、コミット要求に対するアプリケーションへの応答は不明となり、トランザクションの結果はコミットとなり、データベース間の整合性を維持するための動作は何も必要ない。

[0108] 上述してきたように、本実施例では、アプリケーションからのコミット要求に対して関連するディレクタに要求処理部110がコミット処理要求を送信し、応答処理部120がいずれかのディレクタから応答を受け取った場合にアプリケーションにコミットの成功

を通知し、ディレクタがダウンして未解決トランザクションがある場合には、ディレクタが復旧した際に、ディレクタ復旧処理部130が他のディレクタによる未解決トランザクションの処理結果に基づいて未解決トランザクションの処理を復旧したディレクタに指示することとしたので、1フェーズコミットによってデータベース間の整合性を維持し、もってコミット処理を高速化することができる。

[0109] また、ディレクタにコミット処理要求を送信中にコンダクタ100がダウンして一部のディレクタにコミット処理要求が送信できなかった場合には、コンダクタ復旧処理部140が復旧後に各ディレクタに対して未解決トランザクションの問合せを行い、ディレクタが応答した未解決トランザクションに対する他のディレクタの処理結果に基づいて未解決トランザクションの処理を、応答したディレクタに指示することとしたので、1フェーズコミットによってデータベース間の整合性を維持し、もってコミット処理を高速化することができる。

[0110] また、本実施例では、コンダクタ100は、ダウンから復旧した際にディレクタに未解決トランザクションを問合せることによって復旧処理を行うこととしたので、トランザクションログを保有する必要がなく、システムの導入・運用の負担を軽減することができる。

[0111] なお、本実施例では、コンダクタおよびディレクタについて説明したが、コンダクタやディレクタが有する構成をソフトウェアによって実現することで、同様の機能を有するコンダクタプログラムやディレクタプログラムを得ることができる。そこで、これらのコンダクタプログラムやディレクタプログラムを実行するコンピュータについて説明する。

[0112] 図9は、本実施例に係るコンダクタプログラムおよびディレクタプログラムを実行するコンピュータを示す図である。なお、ここでは、コンダクタプログラムとディレクタプログラムを一つのコンピュータで実行する場合について説明するが、コンダクタプログラムとディレクタプログラムを別のコンピュータで実行することもできる。また、アプリケーションとコンダクタプログラムを一つのコンピュータで実行することもできる。

[0113] 同図に示すように、このコンピュータ300は、HDD310と、RAM320と、ROM330と、CPU340と、LANインタフェース350と、I/Oインタフェース360とを有する。

[0114] HDD310は、コンダクタプログラム311およびディレクタプログラム312を記憶する

記憶装置であり、ディレクトプログラム312が使用するデータファイル10、RowIDファイル20およびログファイル30も記憶する。

- [0115] RAM320は、HDD310から読み出されたコンダクタプログラム311およびディレクトプログラム312やプログラムの実行途中結果などを記憶するメモリであり、ROM330は、定数などを記憶した読み出し専用メモリである。
- [0116] CPU340は、RAM320からコンダクタプログラム311およびディレクトプログラム312を読み出し、コンダクタプロセス341およびディレクトプロセス342として実行する処理装置である。
- [0117] LANインタフェース350は、LAN(ネットワーク)と接続するためのインタフェースであり、I/Oインタフェース360は、キーボードやマウスなどの入力装置および表示装置を接続するインタフェースである。
- [0118] そして、このコンピュータ300において実行されるコンダクタプログラム311およびディレクトプログラム312は、フロッピーディスク(FD)、CD-ROM、DVDディスク、光磁気ディスク、ICカードなどの可搬型記憶媒体に記憶され、これらの記憶媒体から読み出されてコンピュータ300にインストールされる。
- [0119] あるいは、コンダクタプログラム311およびディレクトプログラム312は、LANインタフェース350を介して接続された他のコンピュータのデータベースなどに記憶され、データベースなどから読み出されてコンピュータ300にインストールされる。そして、インストールされたコンダクタプログラム311およびディレクトプログラム312は、HDD310に記憶される。

産業上の利用可能性

- [0120] 以上のように、本発明に係る分散トランザクション処理プログラム、分散トランザクション処理方法および分散トランザクション処理装置は、分散トランザクションシステムに有用であり、特に、コミット処理を高速化したい場合に適している。

請求の範囲

- [1] 装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理プログラムであって、
- 前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手順と、
- 前記コミット要求手順により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手順と、
- 前記コミット要求手順により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示手順と、
- をコンピュータに実行させることを特徴とする分散トランザクション処理プログラム。
- [2] 前記復旧指示手順は、前記故障処理装置が復旧したか否かを判定する復旧判定手順と、
- 前記復旧判定手順により故障処理装置が復旧したと判定された場合に、故障処理装置の未解決トランザクションの処理結果について他の処理装置に問合せを行い、問合せ結果に基づいて未解決トランザクションに対する処理を故障処理装置に要求する処理装置復旧処理手順とを、
- コンピュータに実行させることを特徴とする請求項1に記載の分散トランザクション処理プログラム。
- [3] 前記復旧判定手順は、前記故障処理装置に対して定期的に接続を確認することによって該故障処理装置が復旧したか否かを判定することを特徴とする請求項2に記載の分散トランザクション処理プログラム。
- [4] 自装置の故障からの復旧時に処理装置に対して未解決トランザクションを問い合わせる未解決問合せを行い、該未解決問合せに対して処理装置から応答された未解

決トランザクションの処理結果について他の処理装置に問い合わせる処理結果問合せを行う未解決トランザクション処理結果問合せ手順と、

前記未解決トランザクション処理結果問合せ手順による処理結果問合せに対する応答結果に基づいて該未解決トランザクションに対する処理を未解決問合せに対して未解決トランザクションを応答した処理装置に要求する自装置復旧処理手順とを、さらにコンピュータに実行させることを特徴とする請求項1、2または3に記載の分散トランザクション処理プログラム。

- [5] 前記コミット要求手順は、コミット処理の要求がいずれの処理装置にも受け付けられなかった場合に、前記アプリケーションに対してコミットの失敗を応答することを特徴とする請求項1に記載の分散トランザクション処理プログラム。
- [6] 前記コミット応答手順は、前記コミット要求手順により要求したコミット処理に対していずれの処理装置からも応答がなかった場合には、前記アプリケーションにコミット結果を不明と応答することを特徴とする請求項1に記載の分散トランザクション処理プログラム。
- [7] 装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理方法であって、
- 前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求工程と、
- 前記コミット要求工程により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答工程と、
- 前記コミット要求工程により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示工程と、
- を含んだことを特徴とする分散トランザクション処理方法。

- [8] 装置自身の故障以外ではコミット処理要求に対して必ずコミット処理を行って応答する複数の処理装置にアプリケーションから受け付けたトランザクションの分散処理を指示する分散トランザクション処理装置であって、
- 前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手段と、
- 前記コミット要求手段により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手段と、
- 前記コミット要求手段により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示手段と、
- を備えたことを特徴とする分散トランザクション処理装置。
- [9] 一つのトランザクションを分散処理する複数の処理装置と、アプリケーションからトランザクション処理要求を受け付けて該複数の処理装置に分散処理を要求するトランザクション管理装置とから構成される分散トランザクション処理装置であって、
- 前記トランザクション管理装置は、
- 前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求手段と、
- 前記コミット要求手段により要求したコミット処理に対していずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答手段と、
- 前記コミット要求手段により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示手段と、
- を備え、

前記処理装置は、装置自身の故障以外では前記トランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行って該トランザクション管理装置に応答することができるコミット処理手段と、

故障から復旧した際に、前記トランザクション管理装置からの指示に基づいて前記未解決トランザクションの処理を行う復旧処理手段と、

を備えたことを特徴とする分散トランザクション処理装置。

- [10] 一つのトランザクションを分散処理する複数の処理装置と、アプリケーションからトランザクション処理要求を受け付けて該複数の処理装置に分散処理を要求するトランザクション管理装置とから構成される分散トランザクション処理装置のコミット処理方法であって、

前記トランザクション管理装置が前記アプリケーションからのコミット要求に対して複数の処理装置にコミット処理を要求するコミット要求工程と、

前記処理装置が装置自身の故障以外では前記トランザクション管理装置からのコミット処理要求に対して必ずコミット処理を行って該トランザクション管理装置に応答するコミット処理工程と、

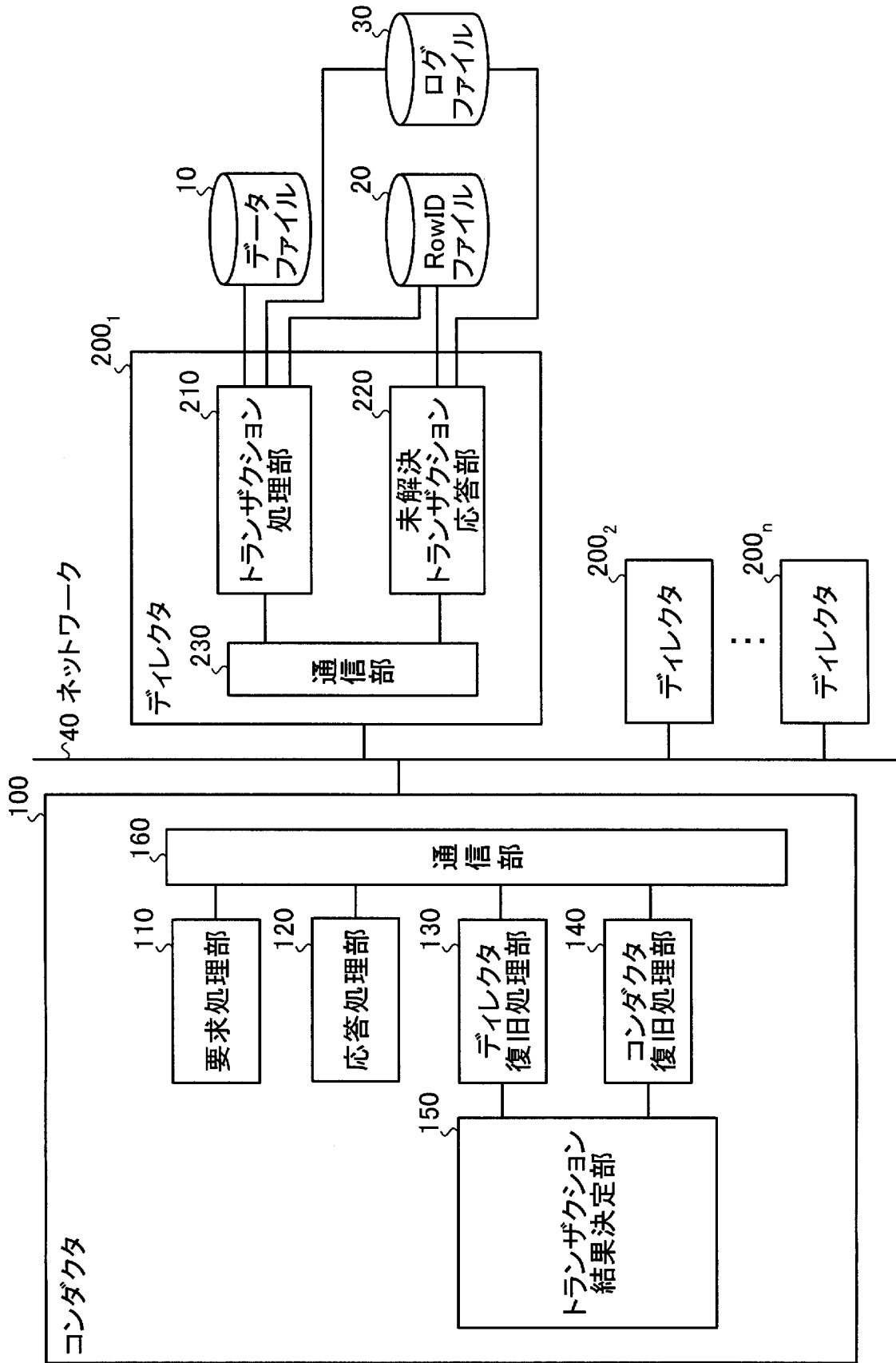
前記コミット要求工程により要求したコミット処理に対して、前記トランザクション管理装置が、いずれかの処理装置から応答があった場合に前記アプリケーションに対してコミットの成功を応答するコミット応答工程と、

前記コミット要求工程により要求したコミット処理に対して故障に起因して応答を行うことができなかった処理装置である故障処理装置が故障から復旧した際に、前記トランザクション管理装置が、該応答が行われなかったトランザクションである未解決トランザクションの処理について他の処理装置の処理結果に基づいて該故障処理装置に対して指示する復旧指示工程と、

故障から復旧した際に、前記故障処理装置が、前記トランザクション管理装置からの指示に基づいて前記未解決トランザクションの処理を行う復旧処理工程と、

を含んだことを特徴とする分散トランザクション処理方法。

[図1]



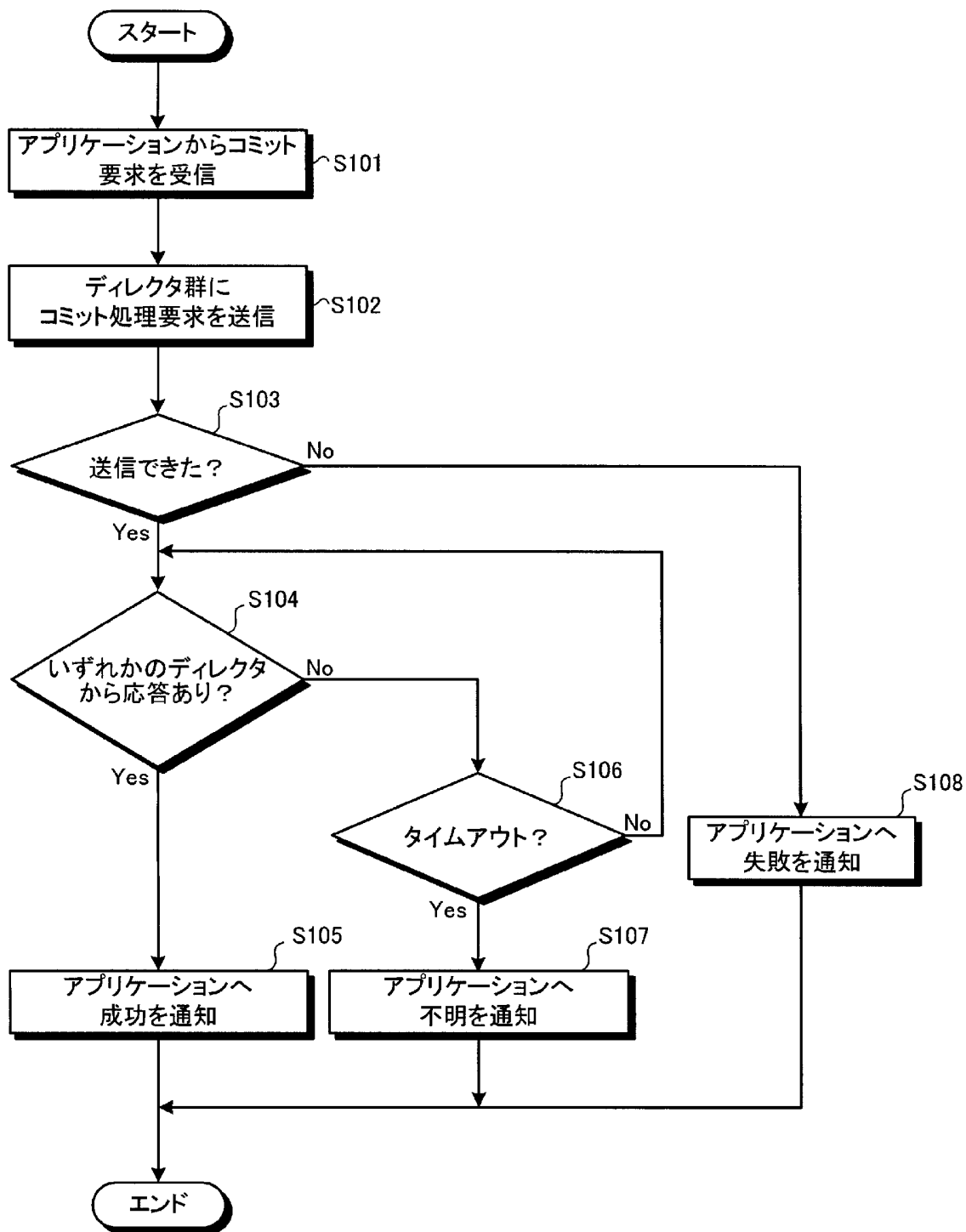
[図2]

レコード格納位置	トランザクションID
----------	------------

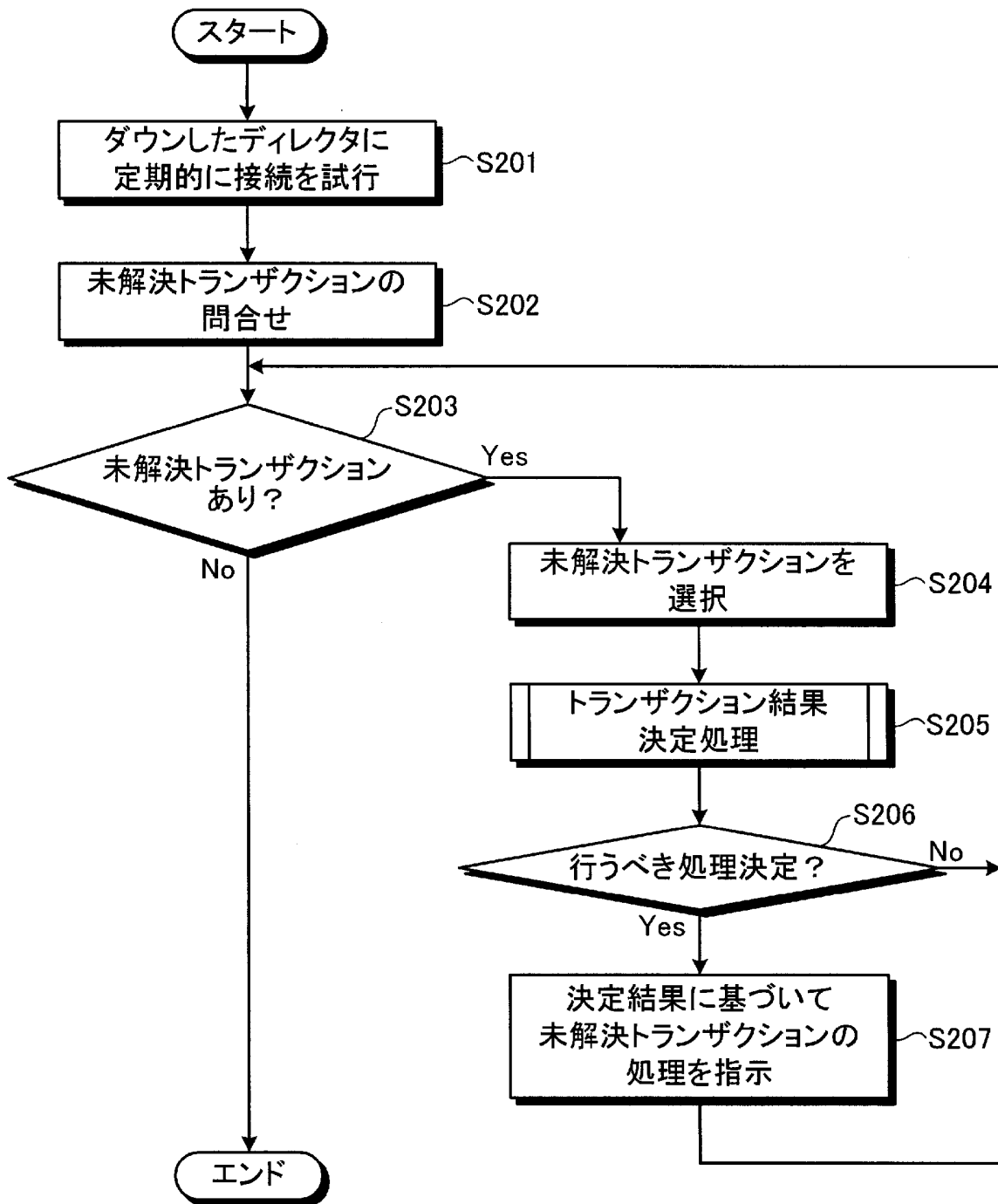
[図3]

トランザクションID ₁
トランザクションID ₂
⋮
トランザクションID _m

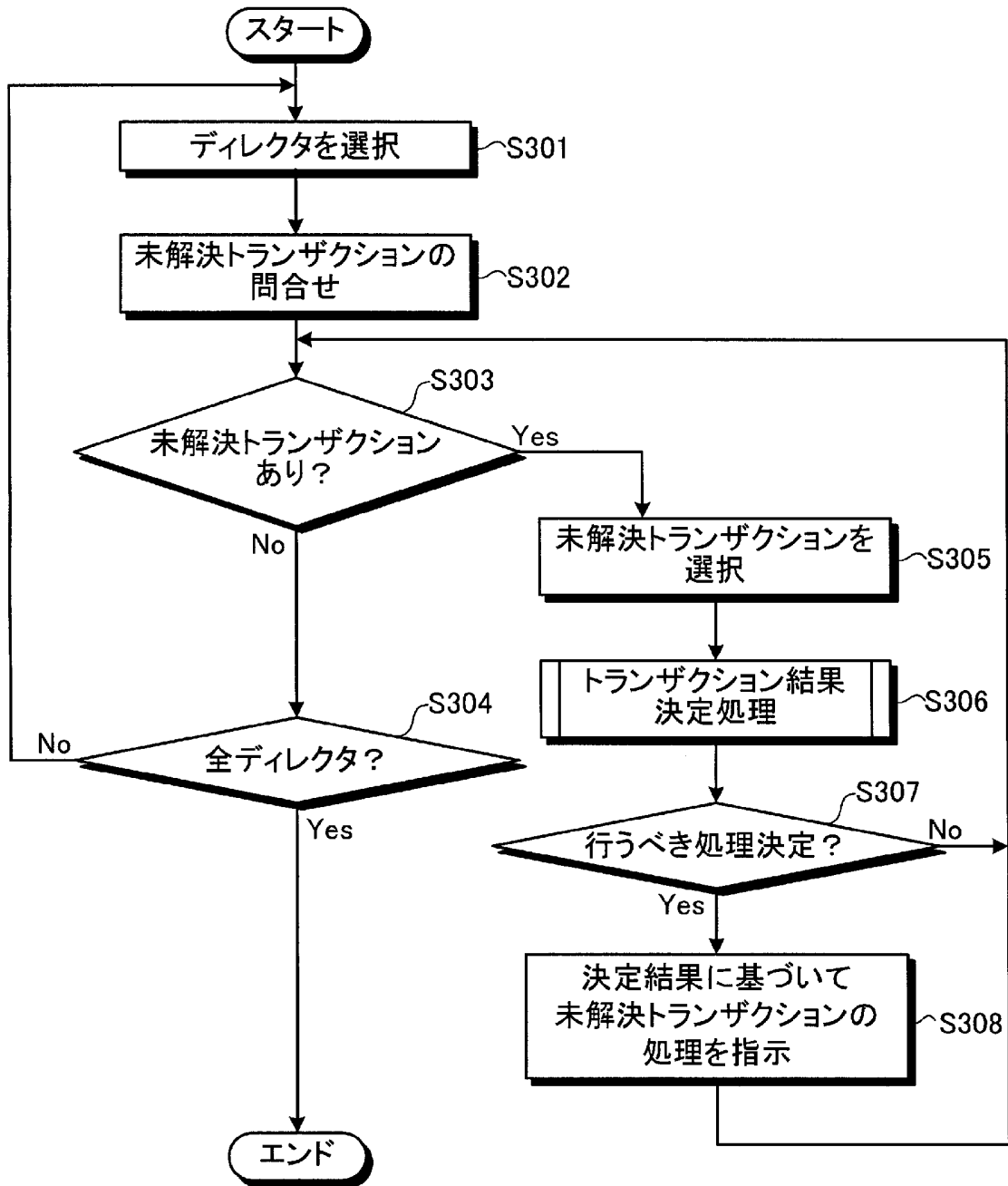
[図4]



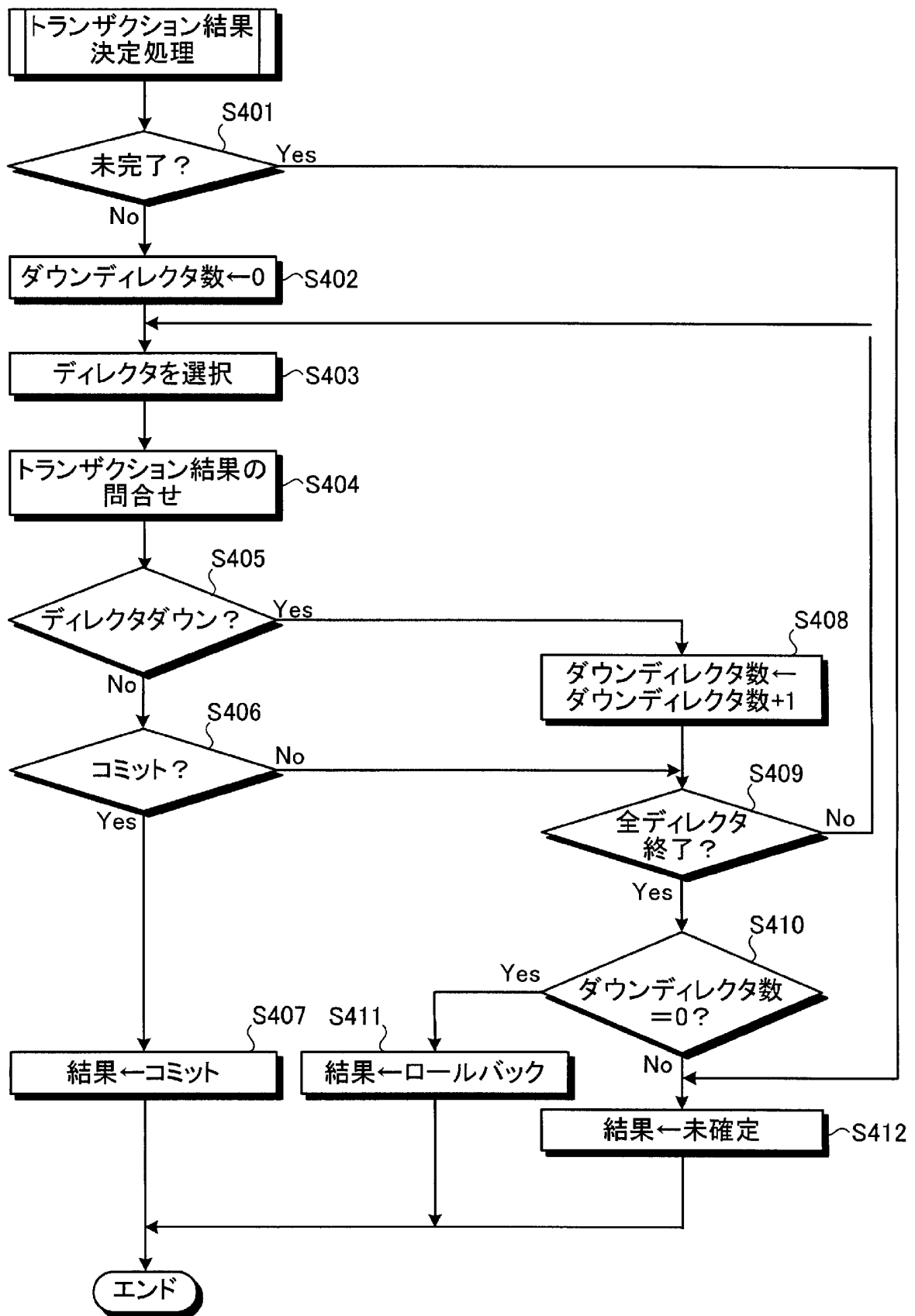
[図5]



[図6]



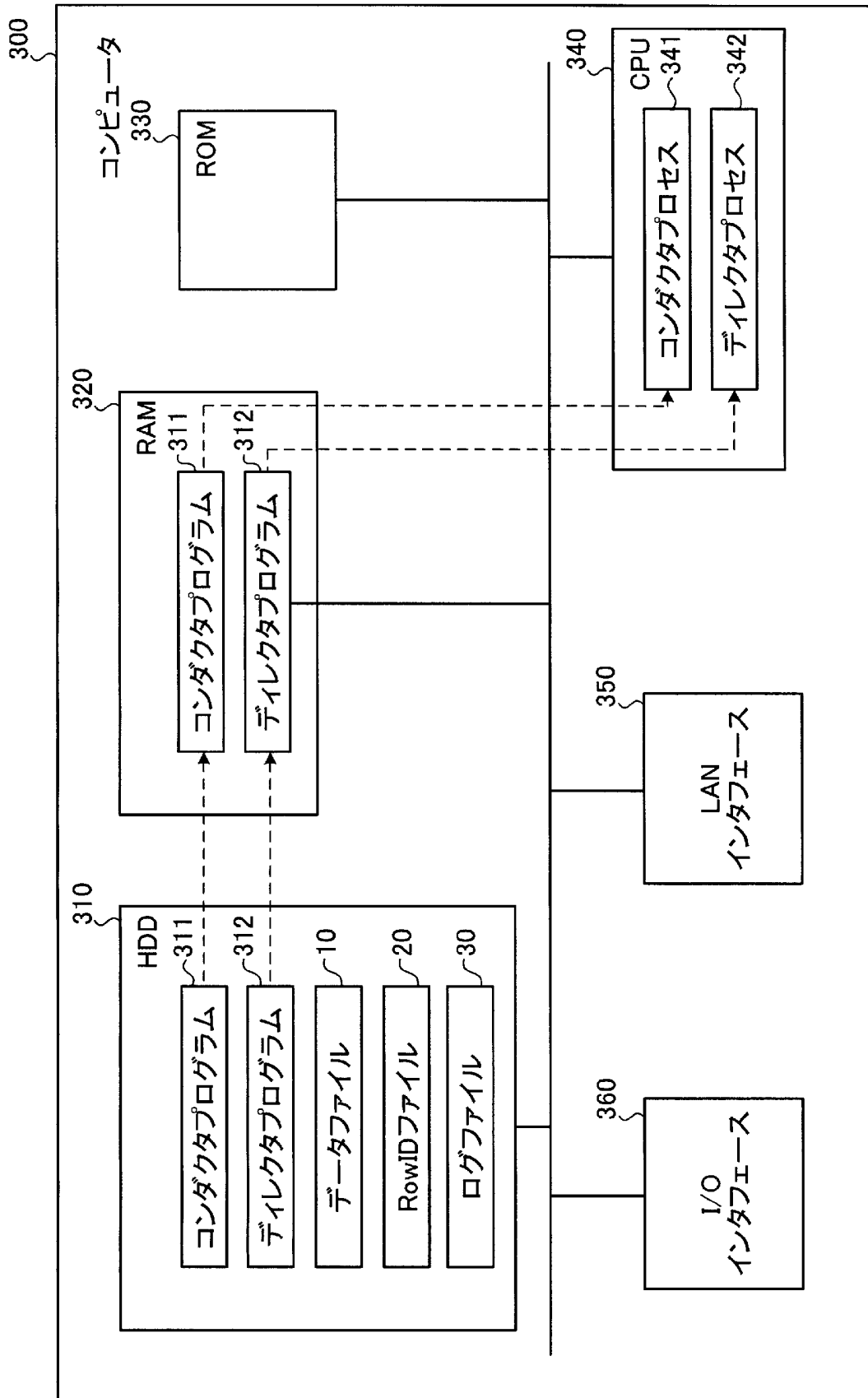
[図7]



[図8]

D ₁ のダウン	D ₂ のダウン	アプリケーションへの 応答	トランザク ションの 結果	整合性維持のための 振る舞い
ダウンしない	ダウンしない	成功	コミット	なし
ダウンしない	コミット指示前	成功	コミット	D ₂ は再起動後、 コンダクタを通じて D ₁ でのコミット完了を 知り、コミットする
ダウンしない	コミット処理中	成功	コミット	D ₂ は再起動後、 コンダクタを通じて D ₁ でのコミット完了を知 り、コミットする
ダウンしない	コミット応答前	成功	コミット	なし
コミット指示前	コミット指示前	失敗	ロールバック	D ₁ 、D ₂ は再起動後、 コンダクタを通じて コミットが完了した ディレクタがなかったこと を知り、 ロールバックする
コミット指示前	コミット処理中	不明	ロールバック	D ₁ 、D ₂ は再起動後、 コンダクタを通じて コミットが完了した ディレクタがなかったこと を知り、 ロールバックする
コミット指示前	コミット応答前	不明	コミット	D ₁ は再起動後、 コンダクタを通じて D ₂ でのコミット完了を 知り、コミットする
コミット処理中	コミット処理中	不明	ロールバック	D ₁ 、D ₂ は再起動後、 コンダクタを通じて コミットが完了した ディレクタがなかったこと を知り、 ロールバックする
コミット処理中	コミット応答前	不明	コミット	D ₁ は再起動後、 コンダクタを通じてD ₂ でのコミット完了を 知り、コミットする
コミット応答前	コミット応答前	不明	コミット	なし

[図9]



(法8条、法施行規則第40、41条)
[PCT18条、PCT規則43、44]

出願人又は代理人 の書類記号 0452284	今後の手続きについては、様式PCT/ISA/220 及び下記5を参照すること。	
国際出願番号 PCT/JP2004/017733	国際出願日 (日.月.年) 29. 11. 2004	優先日 (日.月.年)
出願人 (氏名又は名称) 富士通株式会社		

国際調査機関が作成したこの国際調査報告を法施行規則第41条 (PCT18条) の規定に従い出願人に送付する。
この写しは国際事務局にも送付される。

この国際調査報告は、全部で 3 ページである。

この調査報告に引用された先行技術文献の写しも添付されている。

1. 国際調査報告の基礎

- a. 言語は、下記に示す場合を除くほか、この国際出願がされたものに基づき国際調査を行った。
 この国際調査機関に提出された国際出願の翻訳文に基づき国際調査を行った (PCT規則23.1(b)).

b. この国際出願は、ヌクレオチド又はアミノ酸配列を含んでいる (第I欄参照)。

2. 請求の範囲の一部の調査ができない (第II欄参照)。

3. 発明の単一性が欠如している (第III欄参照)。

4. 発明の名称は 出願人が提出したものを承認する。
 次に示すように国際調査機関が作成した。
分散トランザクション処理方法、装置、及びプログラム

5. 要約は 出願人が提出したものを承認する。
 第IV欄に示されているように、法施行規則第47条 (PCT規則38.2(b)) の規定により国際調査機関が作成した。出願人は、この国際調査報告の発送の日から1カ月以内にこの国際調査機関に意見を提出することができる。

6. 図面に関して

- a. 要約書とともに公表される図は、
第 1 図とする。 出願人が示したとおりである。
 出願人は図を示さなかったため、国際調査機関が選択した。
 本図は発明の特徴を一層よく表しているため、国際調査機関が選択した。
- b. 要約とともに公表される図はない。

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/017733

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G06F12/00

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)

Int.Cl⁷ G06F12/00

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	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

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 4-64146 A (Hitachi, Ltd.), 28 February, 1992 (28.02.92), Full text; all drawings (Family: none)	1-10
A	JP 4-229333 A (International Business Machines Corp.), 18 August, 1992 (18.08.92), Full text; all drawings & US 5319773 A1 & US 5613060 A1 & EP 0457112 A2	1-10

 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

19 August, 2005 (19.08.05)

Date of mailing of the international search report

06 September, 2005 (06.09.05)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/017733

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 10-506742 A (International Business Machines Corp.), 30 June, 1998 (30.06.98), Full text; all drawings & JP 3525933 B2 & US 6202079 B1 & GB 2301686 A & EP 0834122 A1 & WO 1996/039659 A1	1-10
A	JP 7-262073 A (Fujitsu Ltd.), 13 October, 1995 (13.10.95), Full text; all drawings (Family: none)	1-10
A	JP 6-259397 A (Hitachi, Ltd.), 16 September, 1994 (16.09.94), Full text; all drawings (Family: none)	1-10

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. ⁷ G06F12/00										
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. ⁷ G06F12/00										
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 <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2005年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2005年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2005年</td> </tr> </table>			日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2005年	日本国実用新案登録公報	1996-2005年	日本国登録実用新案公報	1994-2005年
日本国実用新案公報	1922-1996年									
日本国公開実用新案公報	1971-2005年									
日本国実用新案登録公報	1996-2005年									
日本国登録実用新案公報	1994-2005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 請求の範囲の番号								
A	JP 4-64146 A (株式会社日立製作所) 1992.02.28, 全文, 全図 (ファミリーなし)	1-10								
A	JP 4-229333 A (インターナショナル・ビジネス・マシーンズ・コーポレーション) 1992.08.18, 全文, 全図 &US 5319773 A1 &US 5613060 A1 &EP 0457112 A2	1-10								
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> 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 19.08.2005	国際調査報告の発送日 06.09.2005									
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) 涌井 智則 電話番号 03-3581-1101 内線 3599	5M 3450								

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 請求の範囲の番号
A	JP 10-506742 A (インターナショナル・ビジネス・マシーンス・コーポレーション) 1998. 06. 30, 全文, 全図 &JP 3525933 B2 &US 6202079 B1 &GB 2301686 A &EP 0834122 A1 &WO 1996/039659 A1	1-10
A	JP 7-262073 A (富士通株式会社) 1995. 10. 13, 全文, 全図 (ファミリーなし)	1-10
A	JP 6-259397 A (株式会社日立製作所) 1994. 09. 16, 全文, 全図 (ファミリーなし)	1-10