

平成 2 6 年 6 月 1 9 日現在

機関番号：4 3 7 0 1

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：2 4 7 0 0 1 0 8

研究課題名（和文）レプリケーション方式の確率モデル適用に関する研究

研究課題名（英文）Studies on Stochastic model of Server Systems with Replication

研究代表者

木村 充位（Kimura, Mitsutaka）

岐阜市立女子短期大学・その他部局等・准教授

研究者番号：2 0 3 5 2 8 2 5

交付決定額（研究期間全体）：（直接経費） 1,300,000 円、（間接経費） 390,000 円

研究成果の概要（和文）：近年，災害等によるデータ損失へ対応し，レプリケーションと呼ばれるネットワークを介して遠隔地にバックアップデータを保存する手法が運用されている．本研究では，ログファイルとデータベースファイルのハイブリッドで一貫性の取れたデータベースの状態に回復できるハイブリッド転送方式や応答時間の短縮と通信負荷の平滑化ができる中継バッファ方式について，マルコフ再生過程を用いた確率モデルにより解析的に分析し，さらに被災による復旧動作やレプリケーションによるコスト変動を解析的に考察してデータベースのバックアップに関する管理・運営について最適方策を提案してきた．成果は国際会議や国際的な学会論文誌に公表してきた．

研究成果の概要（英文）：Recently, the replication has been widely used to protect enterprise database. The backup site stands by the alert when a main site has broken down due to hurricanes, earthquakes, and so on. The server transmits the database content from the main site to the backup site through networks. This is called replication. This research has formulated stochastic models of a server system with replication using a backup method combined database files with journaling files. The other hand, we have formulated stochastic models of a server system with the replication buffering relay method for shortening response time and smooth communication load. Moreover, we have derived the expected costs of the replication and failure recovery analytically. Further, we have discussed optimal policies about database backup method. Research results have been published many international conferences and international journals.

研究分野：総合領域

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：信頼性理論 確率 データベース

1. 研究開始当初の背景

近年、災害によるデータ損失へ対応し、レプリケーションを使ってサーバのデータが保護されている。レプリケーションとは被災に備えてネットワークを介して遠隔地にバックアップデータを保存する手法であり、さらに被災後に業務再開できるサーバを備えたりリモートサイトをもつことによりサーバの早期復旧も実現される。

レプリケーションは、メインサイトのストレージへデータ更新が行われた後、すぐにリモートサイトのストレージへデータ更新を行う同期型や、メインサイトのストレージへデータ更新が行われた後、任意のタイミングでリモートサイトのストレージへデータ更新が行われる非同期型がある(大和純一ら[2006], Oracle Corporation)。

同期型はメインサイトで障害が発生した場合、メインサイトのリスタート時と同じ状況でリモートサイトで業務を再開できる。一方、非同期型はメインサイトで障害が発生した場合、メインサイトでデータ更新完了済みのストレージの内容をリモートサイトで失うことがある。

耐障害システムで用いられる評価尺度として Recovery Point Objective (RPO) 及び Recovery Time Objective (RTO) がある。RPO は、被災によりデータがどのくらい前のものに戻るかを表し、RTO は、被災後システムが再稼動するまでにどのくらいかかるかを表す。RPO は頻繁にバックアップやコピーをとることにより短くすることができる。一方、RTO は遠隔地に現在業務を行っているサイトと同じ装置を近くに揃えておくことができるならば、短くすることができる。従来から、これらの評価尺度に対応すべく、さまざまな方式が提案され議論されている。(今井哲郎ら[2004], 大和純一ら[2004], HiRDB ホワイトペーパー[2004], 中村暢達ら[2007], 渡辺 聡ら[2007])。例えば、データベースアクセスの特性を4種類に別け、その特性ごとに異なるレプリケーション方式を適用するシステムの構築を行った結果、提案方式がネットワーク遅延やパケットロスに対する許容性が従来方式より優れていると評価している(中村暢達ら[2007])。いわばレプリケーションの同期型を採用するか、非同期型を採用するかは、クライアントからのアクセスの特性によって決められる。しかし、これらのモデルではレプリケーションのためのオーバーヘッド増を考慮していない。ここではレプリケーションのためのオーバーヘッド増をコストとして捉え、サーバのダウンに伴うシステムの挙動に確率過程を仮定してレプリケーション方式を評価することを目的とした。

2. 研究の目的

従来までレプリケーションを行うサーバシステムの評価の研究において、データベース更新の同期型や非同期型などのレプリケー

ション方式を適用したサーバシステムの動作検証は、シミュレーションにより解析することが多く、被災による復旧動作やレプリケーションによるさまざまなコスト変動を解析的に考察しているものはない。本研究では、メインサイトとリモートサイトの間に配置した中継装置のバッファにデータを格納することにより、応答時間の短縮と通信負荷の平滑化ができる中継バッファ方式などを代表とする、さまざまなレプリケーション方式(大和純一ら[2004], 菅真樹ら[2005])について、マルコフ再生過程を用いた確率モデル(Osaki[1992])により解析的に分析し、データベースのバックアップに関する管理・運営について最適方策を提案することを目的とした。

3. 研究の方法

(1) 確率モデルの設定と解析

ハイブリッド転送方式、中継バッファ方式など提案されているレプリケーション方式を適用したサーバシステムについてマルコフ再生過程(Osaki[1992].)を用いて確率モデルを構築し、サーバダウンの発生やレプリケーションの時期などに関わるパラメータ特性の相互関連性の調査を行った。その際にモデルの設定方法や解析方法について検証した。また、信頼性解析の研究者らが集まる研究会などで発表を行い議論した。

(2) レプリケーションを導入したサーバシステムに関連するデータの収集

本研究は、理論モデルと実際のサーバにおける数値的差異を少なくするため、データベースを管理運営している担当者との意見交換だけでなく、実際のデータを集めてパラメータにならべく現実に近い値を入力し、プログラムによって数値計算を行い、コスト有効性などを考察した。そのために、実際にサーバシステムにレプリケーションを実装してシミュレーションにより性能評価している文献(根元利弘ら[2002]など)を収集し、その際に得られたクライアントのデータ更新要求の発生分布やそれに伴い発生するジャーナリングファイルの転送分布などを調べ、それらを元に数値計算した。現在までに提案されているレプリケーション方式を適用したサーバシステムについてシミュレーションを行っている論文や書籍、研究報告書などを取り寄せた。

(3) 提案モデルの理論と実際のサーバシステムとの整合性分析

構築してきた基本モデルの改良、修正点を洗い出し、理論モデルの再構築を行った。さらに様々なレプリケーション方式を適用した理論モデルの提案し、これらのモデルの比較検討を軸に研究を進めた。さらに実務面からの意見を取り入れ現実に近いモデルへ近づけるため、サーバ関連など金融機関などでデータベース構築を経験した研究者らが集まる国内外の研究会(学会の全国大会など)や国際会議で発表を行い、モデルの設定方法

や解析方法について知見を広めて、検証した。

4. 研究成果

(1) ハイブリッド転送方式のモデル

概要

近年、提案され実用化されているレプリケーション手法として、データそのものを格納するデータベース(DB)ファイルと、DBの更新履歴情報を時系列に格納したログファイルのハイブリッドで一貫性の取れたデータベースの状態に回復できるハイブリッド転送方式がある(HiRDB ホワイトペーパー[2004])。災害などの発生によりメインサイトのサーバがダウンすると、ダウン直前のリモートサイトにあるDBファイルの状態(チェックポイント)に戻される(ロールバック)が、DBファイル上のデータは中途半端な更新状態となっているため、ログファイル内のDB更新履歴情報を反映し(ロールフォワード)、一貫性のとれた状態に回復する。

本研究では、レプリケーション実行のための閾値として、任意のあるログファイルの転送回数 $k(k=1,2,3,\dots)$ を監視サーバに設定しておくことにより、メインサイトで k 回のログファイル転送が完了後、または一定時間 T の到来後のいずれか早い時期に、直ちにリモートサイトへレプリケーションを実行し、同時にチェックポイント転送を行うモデルを提案し、信頼性に関する諸問題を議論した。

モデルの設定

ここでは、監視サーバとメインサイト及びバックアップサイトで構成される遠隔距離通信のネットワークシステムを考えた。メインサイトとバックアップサイトは、それぞれサーバとストレージで構成されており、バックアップサイトは待機系として常駐する。メインサイトのサーバは通常業務を行い、クライアントの要求によりストレージへデータ更新が行われ、その情報はログファイルとしてバックアップサイトへ同期転送される。また、監視サーバはメインサイトのDBファイルについて、非同期でバックアップサイトのストレージにチェックポイントファイルの転送を含めたレプリケーションの指示を行う。システムの挙動をメインサイト障害発生前のクライアントの要求によるログファイルの転送動作、メインサイトの障害による回復動作、任意のあるログファイルの転送回数 k 回の処理が完了後、または一定時間 T の到来後のレプリケーション動作に着目してマルコフ再生過程を用いて確率モデルで記述した。なお、それぞれの動作の分布は確定できないため、すべて一般分布で記述した。さらに、システムの挙動を示すログファイルの平均転送回数とレプリケーションの平均回数を解析的に求めた。

最適方策の提案

レプリケーションなどの復旧機能の設定と運用に伴って必要となるシステム資源(メモリ空間や時間空間など)の増分をコストとしてとらえ、期待費用を設定した。具体的には1回レプリケーションを行う費用と1回のログファイル転送にかかる費用をある変数で仮定し、(1)の で求めたログファイルの平均転送回数とレプリケーションの平均回数をを用いた総合期待費用を定義して、この総合期待費用を最小にするレプリケーション間隔を解析的に求めた。解析した結果、最適なレプリケーション間隔が存在することを証明できた。

数値的検証

最適なレプリケーション間隔について具体的な数値を求めることによって他のパラメータの変動によって最適なレプリケーション間隔がどのような挙動をするか、数値的な検証を行った。(1)の で仮定した一般分布についてクライアントの要求によるログファイルの転送動作やメインサイトの障害による回復動作はランダムな要因によって発生するため、指数分布を仮定し取り寄せた論文や書籍を参考にしながら数値例を示した。

数値例から、与えられたログファイルの転送回数 k のもとでの最適なレプリケーション間隔 T はデータ更新の平均要求間隔が大きくなるにつれて増加し、メインサイトで障害が発生する平均間隔が大きくなるにつれて増加し、レプリケーションに要する時間が増加するにつれて減少する傾向を示した。与えられた時間 T のもとでの最適なレプリケーション間隔 k も同様な傾向を示した。また、データ更新の平均要求間隔がある程度大きく、レプリケーションの費用がある程度大きい場合は k はレプリケーションに要する時間にほぼ無関係に一定になることがわかった。これは与えられた時間 T でレプリケーションを実行しているために、 k は一定で良いことを示していると考察できる。

(2) 中継バッファ方式のモデル

概要

近年提案され実用化されているバッファ型の中継装置を用いた中継バッファ方式に着目してそのサーバシステムの信頼性の諸問題を議論した。中継バッファ方式とは、クライアントの要求によりメインサイトのストレージへデータ更新が行われたとき、その情報は中継装置へ同期転送される。また、監視サーバは中継装置のDBファイルについて、非同期でバックアップサイトのストレージにレプリケーションの指示を行う方法である(菊池芳秀ら[2004]、菅真樹ら[2005])。しかし、中継装置が保持するデータ更新量が多い場合、メインサイトがダウンした後に中継

装置が保持しているアドレス表の更新・参照に伴うオーバーヘッド増が懸念されるため、どのタイミングで中継装置が保持するデータをバックアップサイトへ転送し、レプリケーションを実行するかは問題である。

本研究では、この問題に対応するため、中継装置の DB ファイルは中継装置において任意のあるデータ更新回数 $n(n=1, 2, 3, \dots)$ を監視サーバに設定しておくことにより、メインサイトで n 回のデータ更新が完了後、直ちにリモートサイトへレプリケーションを実行するモデルを提案し、中継装置が保持しているアドレス表の更新・参照に伴うオーバーヘッド増を考慮した最適なレプリケーション間隔について議論した。

モデルの設定

監視サーバ、中継装置とメインサイト及びバックアップサイトで構成される遠隔距離通信のネットワークシステムを考えた。

メインサイトとバックアップサイトは、それぞれサーバとストレージで構成されており、バックアップサイトは待機系として常駐する。メインサイトのサーバは通常業務を行い、クライアントの要求によりストレージへデータ更新が行われる。また、監視サーバは中継装置の DB ファイルについて、非同期でバックアップサイトのストレージにレプリケーションの指示を行う。

ここでは、メインサイトと中継装置のデータ更新と中継装置からバックアップサイトへのレプリケーションに着目してモデル化を行った。すなわち、監視サーバはメインサイトの状態を常時監視し、更新データを $n(n=1, 2, 3, \dots)$ 回中継装置へ同期転送された後、ネットワークを介して中継装置からバックアップサイトのストレージへ DB ファイルの転送(レプリケーション)を行う。なお、メインサイトでは広域災害が発生した場合、通常業務は一時停止し、バックアップサイトに切り替わり、中継装置とバックアップサイトでシステムが再開始される。クライアントからデータ更新またはデータ読み取り要求が発生し、中継装置またはバックアップサイトでデータ更新またはデータ読み取りが完了する。以上の挙動をマルコフ再生過程を用いて確率モデルで記述した。なお、それぞれの分布はランダムな要因から発生するためすべて指数分布で記述した。システムの挙動を示す中継バッファ装置への平均アクセス回数とレプリケーションの平均回数を解析的に求めた。

最適方策の提案

レプリケーションなどの復旧機能の設定と運用に伴って必要となるシステム資源(メモリ空間や時間空間など)の増分をコストとしてとらえ、期待費用を設定した。具体的には1回レプリケーションを行う費用と中継バッファ装置に格納されているデータに1回

アクセスするのに伴う費用をある変数で仮定し、(2)の で求めた中継バッファ装置への平均アクセス回数とレプリケーションの平均回数を用いた総合期待費用を定義して、この総合期待費用を最小にするレプリケーション間隔を解析的に求めた。解析した結果、クライアントからデータ更新またはデータ読み取り要求が発生したときに、該当する情報が中継装置に格納されている確率を n の関数で仮定した場合に、この関数が n の単調増加であれば最適なレプリケーション間隔が存在することを証明できた。

数値的検証

ここでも最適なレプリケーション間隔について具体的な数値を求めることによって他のパラメータの変動によって最適なレプリケーション間隔がどのような挙動をするか、数値的な検証を行った。(2)の で仮定したクライアントからデータ更新またはデータ読み取り要求が発生したときに、該当する情報が中継装置に格納されている確率を指数分布と仮定し取り寄せた論文や書籍を参考にしながら数値例を示した。

数値例から最適なレプリケーション間隔 n^* はデータ更新の平均要求間隔が大きくなるにつれて減少し、メインサイトで障害が発生する平均発生間隔が大きくなるにつれて増加する傾向を示した。また、レプリケーションに要する時間が増加するにつれて減少する傾向を示した。興味深いことはハイブリッド転送方式では最適なレプリケーション間隔はデータ更新の平均要求間隔が大きくなるにつれて増加していたのに対し、中継バッファ方式では減少する傾向になったことである。これはメインサイトでデータ更新が頻繁に行われる場合は転送処理が複雑になりコストが増加するため中継装置にデータを置いた方が得策であると考察できる。

(3) 得られた成果の国内外における位置づけとインパクト

サーバシステムにハイブリッド転送方式や中継バッファ方式を用いてデータバックアップを行った場合、ログファイルの転送のオーバーヘッド増やレプリケーションのオーバーヘッド増を考慮してバックアップする時期を解析的に検討することは意義あることであり、有効なレプリケーション手法であるハイブリッド転送方式や中継バッファ方式をより有効的な方法で運用する最適方策を確率モデルにより一般化できたことはサーバシステムの効率的な運用法に波及できるのではないかと考える。

当初は方式間の比較を考えていたが、それぞれに一長一短なところがあり、これらの方式を組み合わせたモデルを考えて検証する予定である。実際のシステムを用いて検証はできなかったものの、(1)や(2)のようにモデルを一般化できたことは、実際に具体

的な分布が仮定できた場合、現実のデータベースのバックアップに関する管理・運営について有効な手段になりえると考えます。

このような広域災害等に対応したサーバシステムの高信頼化の問題は、今後ますます重要な課題となることが考えられ、この方面に対する多くの研究が期待される。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計6件)

Mitsutaka Kimura, Mitsuhiro Imaizumi, Toshio Nakagawa, Reliability Analysis of a Server System with Replication Buffering Relay Method, International Journal of Reliability and Quality Performance, 査読有、印刷中

Mitsutaka Kimura, Mitsuhiro Imaizumi, Toshio Nakagawa, RELIABILITY MODELING OF A SERVER SYSTEM WITH HYBRID REPLICATION METHOD, Proceedings of Nineteenth ISSAT International Conference on Reliability and Quality in design, 査読有、vol.19、2013、pp.43-47

Mitsutaka Kimura, Mitsuhiro Imaizumi, Toshio Nakagawa, RELIABILITY ANALYSIS OF A REPLICATION WITH LIMITED NUMBER OF JOURNALING FILES, Reliability Engineering & System Safety, 査読有、vol.116、2013、pp.105-108

DOI: 10.1016/j.ress.2013.02.008

Mitsutaka Kimura, Mitsuhiro Imaizumi, Toshio Nakagawa, Optimal Policy of a Server System with Replication Buffering Relay Method, Proceeding of 2012 Workshop on Dependable Transportation Systems/Recent Advances in Software Dependability, 査読有、vol.18、2012、pp.45-49

Mitsutaka Kimura, Mitsuhiro Imaizumi, Toshio Nakagawa, RELIABILITY MODELING FOR A SERVER SYSTEM WITH BUFFERING RELAY METHOD, Proceeding of Advanced Reliability and Maintenance Modeling V, 査読有、vol.5、2012、pp.255-262

〔学会発表〕(計4件)

木村充位、レプリケーション機能をもつサーバシステムの信頼性的考察、日本オペレーションズ・リサーチ学会、第41回中部支部研究発表会、2014年03月15日～2014年03月15日、ウインクあいち(愛知県)

木村充位、中継バッファ方式と同期レプリケーションを併用したサーバシステムの最適方策、日本オペレーションズ・リサーチ学会、2014年春季研究発表会、2014年03月06日～2014年03月07日、大阪大学豊中キャンパス(大阪府)

木村充位、ディザスタリカバリシステムにおけるハイブリッド転送方式の最適方策、日本オペレーションズ・リサーチ学会、2014年春季研究発表会、2013年03月05日～2013年03月06日、東京大学(東京都)

木村充位、ディザスタリカバリシステムにおける中継バッファ方式の評価方法、日本オペレーションズ・リサーチ学会(招待講演)、2012年09月12日～2012年09月13日、ウインクあいち(愛知県)

〔図書〕(計1件)

Mitsutaka Kimura、World Scientific、Reliability Modeling with Applications Essays in Honor of Professor Toshio Nakagawa on His 70th Birthday (edited by Syouji Nakamura)、2014、pp.157-176

〔産業財産権〕

出願状況(計0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

<http://dandelion.gifu-cwc.ac.jp/wp-content/themes/gcwc/img/page/pdf/kimura.pdf>

6. 研究組織

(1) 研究代表者

木村 充位 (KIMURA Mitsutaka)
岐阜市立女子短期大学・国際文化学科・
准教授
研究者番号：20352825

(2) 研究分担者

()
研究者番号：

(3) 連携研究者

()
研究者番号：