

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4025650号
(P4025650)

(45) 発行日 平成19年12月26日(2007.12.26)

(24) 登録日 平成19年10月12日(2007.10.12)

(51) Int.Cl.

F I

G06F 11/00 (2006.01)

G06F 9/06 630C

請求項の数 9 (全 8 頁)

(21) 出願番号	特願2002-573368 (P2002-573368)	(73) 特許権者	503148270
(86) (22) 出願日	平成14年3月12日(2002.3.12)		テレコム・イタリア・エッセ・ビー・アー
(65) 公表番号	特表2004-526246 (P2004-526246A)		イタリア国 アイー20123 ミラノ
(43) 公表日	平成16年8月26日(2004.8.26)		ピアッツァ デグリア アッファーリ 2
(86) 国際出願番号	PCT/IT2002/000152	(74) 代理人	100064355
(87) 国際公開番号	W02002/076020		弁理士 川原田 一穂
(87) 国際公開日	平成14年9月26日(2002.9.26)	(72) 発明者	ジャンニ・カナル
審査請求日	平成17年3月4日(2005.3.4)		イタリア国 アイー10148 トリノ
(31) 優先権主張番号	T02001A000264		ヴィア ジー・ レイス ロモリ 274
(32) 優先日	平成13年3月21日(2001.3.21)		、テレコム・イタリア・エッセ・ビー・ア
(33) 優先権主張国	イタリア(IT)		ー内

最終頁に続く

(54) 【発明の名称】 ネットワークサーバーのプログラムを更新する方法、その関連システム及びソフトウェア

(57) 【特許請求の範囲】

【請求項1】

少なくとも1つの現在のプロセスプログラム(1)を実行するネットワークサーバー(S)のプログラムをアップグレードする方法であって、前記現在のプロセスプログラム(1)は、ネットワークパケットとインターフェースする現在の下位レイヤ(2)と前記パケットを処理するための現在の上位レイヤ(3)とを含む前記方法において、

ー ネットワークパケットとインターフェースする更新バージョンの下位レイヤ(2')と前記パケットを処理するための更新バージョンの上位レイヤ(3')とを含んだ更新バージョンのプロセスプログラム(1')を付与するステップであって、前記更新バージョンの下位レイヤ(2')は、通信チャネル(4)に接続されてネットワークパケットを前記更新バージョンの下位レイヤ(2')から前記現在の上位レイヤ(3)に迂回させる前記ステップ、

ー 前記現在の下位レイヤ(2)を前記更新バージョンの下位レイヤ(2')と置換するステップ(106)、

ー 前記更新バージョンの下位レイヤ(2')から受け取ったパケットが前記現在の上位レイヤ(3)にて既に進行中のトランザクション又は新しいトランザクションのどちらに關係するかを検査するステップ(112)であって、この検査ステップ(112)を前記現在の上位レイヤ(3)で実行する前記ステップ、

ー 前記現在の上位レイヤ(3)にて既に進行中のトランザクションに関するパケットを前記現在の上位レイヤ(3)において処理するステップ(114)、

10

20

ー 新しいトランザクションに関するパケットを前記更新バージョンの上位レイヤ（３’）において処理するステップ（１１６）、
ー 前記現在の上位レイヤ（３）にて既に進行中のトランザクションが完了したかどうかを判断するステップ（１１８）、及び
ー 前記現在の上位レイヤ（３）にて既に進行中の前記トランザクションが完了すると、前記通信チャンネル（４）を閉鎖し、サーバー（Ｓ）のプログラムアップグレードを終了するステップ（１２０）を含むことを特徴とする前記方法。

【請求項２】

前記プログラムアップグレードを実行するために、該プログラムアップグレードが単一のサーバー（Ｓ）にて実行されることを検査するステップ（１０２）を含むことを特徴とする請求項１に記載の方法。

10

【請求項３】

前記現在の下位レイヤ（２）を前記更新バージョンの下位レイヤ（２’）と置換する前記ステップが、前記現在の下位レイヤ（２）のネットワークポートを前記更新バージョンの下位レイヤ（２’）に移転するステップを含むことを特徴とする請求項１又は２に記載の方法。

【請求項４】

前記現在の上位レイヤ（３）が、前記通信チャンネル（４）によって新しいトランザクションに関するパケットを前記更新バージョンの上位レイヤ（３’）に送るように構成されていることを特徴とする請求項１～３のいずれか一項に記載の方法。

20

【請求項５】

アップグレードされる前記現在のプロセスプログラム（１）とアップグレードされたプロセスプログラム（１’）の間で状態情報を交換する（１０８）ように前記通信チャンネル（４）を構成するステップを含むことを特徴とする請求項１～４のいずれか一項に記載の方法。

【請求項６】

前記通信チャンネル（４）をソケットとして構成するステップを含むことを特徴とする請求項１～５のいずれか一項に記載の方法。

【請求項７】

請求項１～６のいずれか一項に記載の方法を用いてプログラムをアップグレードするように構成されたネットワークサーバー。

30

【請求項８】

前記サーバーがセッション・イニシエーション・プロトコル（ＳＩＰ）プロキシであることを特徴とする請求項７に記載のサーバー。

【請求項９】

コンピュータメモリに直接ロードすることができるソフトウェアであって、コンピュータ上にて実行する際に請求項１～６のいずれか一項に記載の方法を実施するプログラムコード部分を含んだ前記ソフトウェア。

【発明の詳細な説明】

40

【技術分野】

【０００１】

本発明は、サーバーの使用を伴うネットワークに関し、特に、本発明は、これらのサーバーのプログラムを更新するための機能、言い換えれば、現在「ソフトウェアアップグレード」と称されている機能に関する。特に、本発明による方法は、トランザクションを管理し且つ状態（もしそれを損失すれば永続的又は一時的な故障を引き起こし得る）を保持するサーバーにおける潜在的な用途に特に留意しつつ開発された。

【背景技術】

【０００２】

サービスを中断せず又は最小の中断回数のみで実行するのを保証する必要がある場合、

50

サービスを利用できないことが生じ得るようなサーバーソフトウェアアップグレードは特に注意が必要となる。

【0003】

この点について、ある程度のフォールト・トレランスを保証するために一般に用いられる解決策は、サーバーの複製である。更新オペレーションが実施されると、このオペレーションに関係するサーバー（単数又は複数）が一時的に分離される。この分離は、一般にはこれらのサーバーを探し出すことができるデータベースを例えばドメイン・ネーム・サーバー（DNS）機構により削除し、使用中のサーバーに負荷を移すことにより行われる。

【0004】

この解決策の主な欠点は次の通りである。

10

（１）いつサーバーを動作停止できるかを判断するのが常に容易なわけではないこと。すなわち、実際には、進行中のトランザクションが存在していないこと、及びクライアントとして動作している装置におけるサーバーリストのローカルコピーが当該サーバーを参照していないことを保証する必要がある。

（２）この欠点に直接関係する別の問題は、管理オペレーションが常に相当遅く、特に緊急な場合（例えば起こりうるセキュリティ問題に対処する必要がある場合）唯一できるのは、サービスを一時的に中断することである。

（３）加えて、該プロセスにより維持される状態の所要時間がトランザクションの所要時間を超える場合（しばしばそうであるように）、その有効な生存期間の終わりを容易に識別することができず、その状態を喪失し一時的な故障を引き起こしてしまうことである。

20

（４）最後に、（たいていの場合にそうであるように）トラヒック傾向を予想するのが難しく、更新オペレーション中にトラヒックピークが生じるならば、動作中のサーバーが、それらの作業負荷を処理することの難しさに直面し得ることである。この問題を防止する唯一の方法は、通常の状態下で十分に対応できる機器に加えて少なくとも１つの機器を配備することであり、このことにより、厳密には必要ないシステム冗長性を増大させる。

【発明の開示】

【発明が解決しようとする課題】

【0005】

従って、本発明の目的は、上記概説した問題を解消することができる解決策を提供することである。

30

【課題を解決するための手段】

【0006】

本発明により、この目的は、特許請求の範囲に詳細に記した特徴を備えた方法により達成される。本発明はまた、それに関連するシステム、言い換えれば、前述の方法に基づいて動作するよう構成されたシステムに関し、また、それに対応したソフトウェア製品、言い換えれば、コンピュータにロードされ該コンピュータ上で実行されるコードであって、該コンピュータが前述の方法に基づいて動作できるように該コンピュータを構成することができるコードにも関する。

【0007】

基本的に、本発明による解決策は、「ホット」サーバープロセス置換を伴い、置換されるプロセス（被置換プロセス）がその置換操作と一定期間共存するように進行する。このオペレーションの所要時間は、被置換プロセスにより決められ、よって常に該オペレーションが最後のトランザクションの終了後に行われることが保証される。より長い生存期間を有する状態情報は、例えばソケット又は別の双方向性プロセス間通信機構により一方のプロセスから他方のプロセスに送ることができる。

40

【0008】

図面の簡単な説明

以下、添付図面を参照して本発明を説明するが、単に例として行っており、限定的に解釈すべきでない。

図１は、本発明に従って動作するサーバーの通常オペレーションを示す機能ブロック図

50

であり、

図2は、図1と同じ形式で示した別の機能ブロック図であり、図1に示したサーバー更新オペレーションをどのように実行するかを示しており、

図3は、図2に示したオペレーションをどのように実行するかを示したフローチャートである。

【発明を実施するための最良の形態】

【0009】

図1は、サーバーSの一般的な構成の略図である。このサーバーSは、例えば、通信ネットワークにおいて用いられてSIP（セッション・イニシエーション・プロトコル）プロキシ機能を実行するサーバーである。

10

【0010】

これは、例えば、Solaris（登録商標）オペレーティングシステムを用いる、Sun Microsystems (Palo Alto (CA)) により製造されたEnterpriseサーバーである。

【0011】

サーバーSは、その全体が1で示されたようなプロセスオペレーションを実行することができる。

【0012】

オペレーション1は、次のものを含む。

(1) パケットの受信／送信を保証するためにネットワークと直接インターフェースを行う下位レイヤ（レベル）2（以下ではレイヤとレベルなる用語は等価なものとする）、及び

20

(2) パケットを受け取り、処理し、下位レイヤに送る上位レイヤ3。

【0013】

2つのレイヤ2及び3は、矢印で表されているような一般的な双方向通信方式にて接続される。

【0014】

上述した点すべてが、当業者に十分よく知られることになるであろう構成及び適用の基準に対応していることが分かるので、ここで詳細な説明をする必要はないであろう。

【0015】

サーバーのプログラムの更新（言い換えれば、いわゆるソフトウェア・アップグレード・オペレーション）を行う際、1で示した現在の構成を、全体を1'で示した更新構成と置換しなければならず、この更新構成も同様に下位レイヤ2'と上位レイヤ3'に編成されており、更新オペレーションすなわちアップグレードオペレーションの終わりに対応するレイヤ2及び3を取り替える。

30

【0016】

本発明による方法は、現在のプロセス1をその更新バージョン1'に取り替え又は置換するオペレーションがネットワークとは完全に無関係であることを保証するものであり、言い換えれば、サーバーSに接続されたクライアントがアップグレードオペレーションを開始したこと及びこのオペレーションが進行中であることを実際には検出できないように該オペレーションが実施されることを保証する。

40

【0017】

これらのことすべてが、該プロセスにより維持される状態の所要時間がトランザクションの所要時間を超える（それにより、有効な生存期間の終わりをある程度容易に識別するのがしばしば不可能となる）場合でさえ、進行中のトランザクションが正常に管理されることを保証し、かつ予想できないトラヒックピークに続いて生じ得るどんな困難も回避しつつ実現される。

【0018】

アップグレードオペレーションでは、まず、更新又はアップグレードに対応するプロセス1'が、全体として4で示され自身を認証する(authenticating itself) ソケット又は他のIPC機構によって、置換すべき現在のプロセス1に接続される。

50

【0019】

当然、置換される現在のプロセス1は、アップグレードオペレーションが同一の機器Sにより実行されることの検査も行う。

【0020】

置換される現在のプロセス1は、そのポート（下位レイヤ2）を解放する。これらのポートは、アップグレードされたプロセス1'の対応するレイヤ2'によりすぐを取得される。このレイヤ2'は、受信したパケットフローをチャネル4を介して現在のプロセス1の上位レイヤ3に送るために、このパケットフローを選択的に迂回させることができるように構成されている。

【0021】

それから、ネットワークから到着したパケットは、アップグレードされたプロセス1'の下位レイヤ2'により専ら受信される。

【0022】

いずれにしても、これらのパケットは、そのとき置換されつつあるプロセス1の上位レイヤ3にも送られる。よって、このプロセス1の上位レイヤ3は、いわばフィルターとして作用する立場に置かれる。すなわち、もしメッセージが新しいトランザクション（言い換えれば、アップグレードオペレーションの開始後に始まったトランザクション）に関係するならば、アップグレードされたプロセス1'の対応する上位レイヤ3'に対して、処理することなく直接メッセージを転送する。逆に、もしメッセージが、置換されつつあるプロセス1の上位レイヤ3にて既に進行中のトランザクションに関係するならば、この上位レイヤ3によりメッセージを続けて処理する。

【0023】

加えて、2つのプロセス1及び1'は、トランザクションに関わりのない状態情報を交換することができる。

【0024】

置換されつつあるプロセス1の上位レイヤ3は、それが管理するトランザクションのすべてが終了したことを判断し、アップグレードオペレーションが完了したことを示す信号を発生する。この信号は、チャネル4上に送られ、アップグレードされたプロセス1'の下位レイヤ2'が、アップグレードされたプロセス1の上位レイヤ3に対してどんな種類のパケット又はメッセージもさらに送らないことを保証する。通信チャネル4は閉鎖され、よって、アップグレードされたプロセス1'がこの時点であらゆる点でプロセス1を置換しており、プロセス1は停止される。

【0025】

上述したアップグレードオペレーションを実行するのに用いられるステップを、図3のフローチャートにさらに詳細に示す。

【0026】

このフローチャートにおいて、符号100はアップグレードオペレーションにおける最初のステップを示し、これは接続4の開設に対応する。

【0027】

符号102は、同一の機器にアップグレードが実行されていることを現在のプロセス1が検査するステップを示す。

【0028】

この検査の結果が否定（結果F）であるとき、アップグレードオペレーションを継続せず、ステップ104にて中断する。

【0029】

ステップ102で実行する検査の結果が肯定（結果V）であるとき、ステップ106においてシステムポート（実際にはネットワークとのインターフェース）を現在のプロセス1の下位レイヤ2からアップグレードされたプロセス1'の下位レイヤ2'に移し、次にステップ108において状態情報を転送する。

【0030】

10

20

30

40

50

この時点以降、既に進行中のトランザクションに関するメッセージは現在のプロセス1により取り扱われ続け一方、新しいトランザクションに関するメッセージはアップグレードされたプロセス1'により取り扱われるという動作条件が設定される。

【0031】

特に、符号110は、新しいメッセージを読み取るステップを示し、符号112は、ちょうど読み取られたパケットが既知のトランザクションと新しいトランザクションのどちらに関係するかを判断する検査を実行するステップを示す。

【0032】

メッセージが新しいトランザクションに関係する場合（ステップ112の結果がF）、パケットは「新しい」トランザクションに関するものと識別され、よってステップ114 10
においてアップグレードされたプロセス1'により取り扱われる。

【0033】

読み取られたメッセージが、既知のトランザクションに対応するものと識別された場合（ステップ112の結果がV）、そのメッセージはステップ116において現在のプロセス1により取り扱われる。

【0034】

ステップ114とステップ116の両方から開始して、アップグレードオペレーションは継続し、118で示した別のステップに合流し、このステップ118において、置換されつつあるプロセス、言い換えれば現在のプロセス1に対するトランザクションが終了したと見なせるか否かを判断する。 20

【0035】

まだ完了していないトランザクションがなお存在している場合（ステップ118の結果がF）、新しいパケットを読み取るならば、システムオペレーションは、ステップ110の上流に再び戻る。

【0036】

逆に、もし置換されつつあるプロセス1により取り扱う必要のあるトランザクションが終了したならば（ステップ118の結果がV）、システムは120で示される最終ステップに進み、このステップ120において通信チャンネル4を閉鎖し、よってアップグレードオペレーションを完了する。

【0037】

当然、本発明の基礎となる原理を損なうことなく、ここでの記載及び説明に関して本発明の範囲を逸脱することなく実施の詳細及び形態を広く変えることができる。 30

【0038】

例えば、図1及び3に関して上述した一般的なアプローチにつき、単なる例示であってあらゆる点で限定的に解釈すべきでない可能な変更として、次のものを挙げることができる。

（1）チャンネル4により示される通信機構がソケットでない。すなわち、本質的なことは、一般にこのチャンネルにより双方向の通信ができること、及び同様に一方のプロセスが実行されつつある場合に確実に設定できることである。

（2）下位レイヤ2を下位レイヤ2'と取り換え又は置換する初期機構が、前述のポート 40
に加えて又はこれらのポートの代わりに、専ら入力／出力機能に用いられる他の専用リソースに適用される。

【0039】

原則的に、本発明による方法は、サービスの中断時間を最小（実質的には零）にしなければならず、且つ取り扱われるトラフィックがかなり少ない期間を識別することができないというあらゆる状況において適用される。

【図面の簡単な説明】

【0040】

【図1】本発明に従って動作するサーバーの通常オペレーションを示す機能ブロック図である。 50

【図 2】図 1 と同じ形式で示した別の機能ブロック図であり、図 1 に示したサーバー更新オペレーションをどのように実行するかを示す。

【図 3】図 2 に示したオペレーションをどのように実行するかを示したフローチャートである。

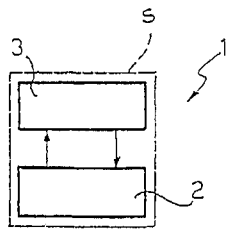
【符号の説明】

【0041】

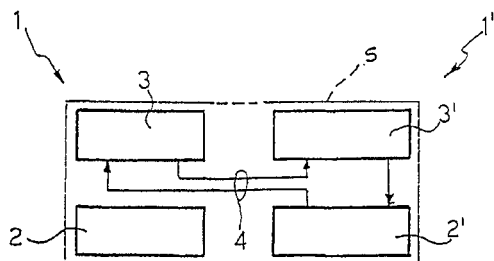
- 1 被置換プロセス
- 1' 置換プロセス
- 2 下位レイヤ
- 3 上位レイヤ
- 4 通信チャンネル
- S サーバー

10

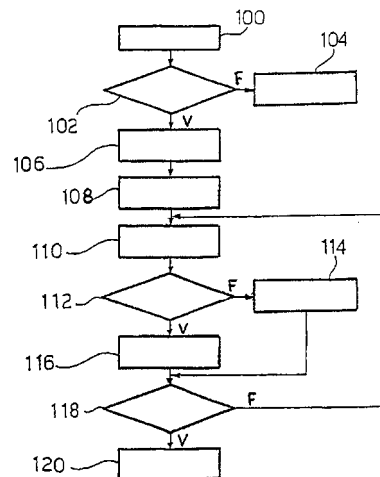
【図 1】



【図 2】



【図 3】



フロントページの続き

(72)発明者 アルベルト・クダ
イタリア国 アイー１０１４８ トリノ ヴィア ジー、 レイス ロモリ ２７４、テレコム・
イタリア・エッセ・ピー・アー内

審査官 漆原 孝治

(56)参考文献 特開平０６－３１９１６３（ＪＰ，Ａ）

(58)調査した分野(Int.Cl., DB名)
G06F 11/00