

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5042318号  
(P5042318)

(45) 発行日 平成24年10月3日(2012.10.3)

(24) 登録日 平成24年7月20日(2012.7.20)

(51) Int.Cl.

G06F 13/00 (2006.01)  
H04L 7/00 (2006.01)

F 1

G06F 13/00  
H04L 7/00351C  
Z

請求項の数 12 (全 41 頁)

(21) 出願番号 特願2009-547641 (P2009-547641)  
 (86) (22) 出願日 平成20年1月23日 (2008.1.23)  
 (65) 公表番号 特表2010-528345 (P2010-528345A)  
 (43) 公表日 平成22年8月19日 (2010.8.19)  
 (86) 国際出願番号 PCT/EP2008/050739  
 (87) 国際公開番号 WO2008/092781  
 (87) 国際公開日 平成20年8月7日 (2008.8.7)  
 審査請求日 平成22年9月22日 (2010.9.22)  
 (31) 優先権主張番号 60/887,562  
 (32) 優先日 平成19年1月31日 (2007.1.31)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 390009531  
 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  
 INTERNATIONAL BUSINESS MACHINES CORPORATION  
 アメリカ合衆国10504 ニューヨーク  
 州 アーモンク ニュー オーチャード  
 ロード  
 (74) 代理人 100108501  
 弁理士 上野 剛史  
 (74) 代理人 100112690  
 弁理士 太佐 種一  
 (74) 代理人 100091568  
 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】協定タイミング・ネットワーク内の階層1構成を定義する方法、システム、およびコンピュータ・プログラム

## (57) 【特許請求の範囲】

## 【請求項1】

協定タイミング・ネットワークの1つまたは複数のサーバのうちの第1のサーバにより、前記協定タイミング・ネットワークの構成に関する情報を入手するステップであって、前記協定タイミング・ネットワークがクロック同期用のタイミング機能を有し、前記協定タイミング・ネットワークにより任意の時点で1つのサーバのみが階層1サーバになることができ、階層1サーバが、前記協定タイミング・ネットワークの前記1つまたは複数のサーバのすべての同期のためにルート・クロック・ソースを提供することを特徴とし、

前記情報が、前記1つまたは複数のサーバのうちのどのサーバが階層1サーバの役割を引き受けることができるかを示す情報を含み、前記構成が複数のタイプの構成のうちの1つのタイプの構成であるステップと、

前記サーバにより、前記入手した情報に基づいて、前記協定タイミング・ネットワーク内の前記第1のサーバの役割を判断するステップと、

前記役割が階層1サーバであると判断したことに応答して、前記1つまたは複数のサーバの同期のためにルート・クロック・ソースを提供するステップと、

前記役割が階層1サーバではないと判断したことに応答して、前記階層1サーバからルート・クロック同期を受信するステップとを含み、

前記複数のタイプの構成が、マル構成と、単一サーバ構成と、デュアル・サーバ構成と、トライアッド構成とを含み、

1次アクティブ・サーバが未定義であるマル構成に前記サーバがなっていることを前記

情報が示し、その役割が非同期サーバであることを前記サーバが判断するステップと、前記サーバが、非ヌル階層1構成を有する他のサーバに接続することをさらに含み、接続したことに応答して、前記サーバを前記協定タイミング・ネットワーク内で同期状態にするステップをさらに有する、

方法。

【請求項2】

1次アクティブ・サーバが定義されている単一サーバ構成に前記サーバがなっていることを前記情報が示し、その役割が前記1次アクティブ・サーバまたは前記1次アクティブ・サーバからタイミング情報を受信する2次サーバのいずれかであることを前記サーバが判断する、請求項1記載の方法。

10

【請求項3】

前記判断が、前記サーバに維持された階層1構成情報ブロック内の情報に基づいて行われる、請求項2記載の方法。

【請求項4】

1次アクティブ・サーバおよび代替サーバが定義されているデュアル・サーバ構成に前記サーバがなっていることを前記情報が示し、その役割が、前記1次アクティブ・サーバ、前記代替サーバ、または前記1次アクティブ・サーバあるいは前記代替サーバのいずれでもないサーバであることを前記サーバが判断する、請求項1記載の方法。

【請求項5】

前記判断が、前記サーバに維持された階層1構成情報ブロック内の情報に基づいて行われる、請求項4記載の方法。

20

【請求項6】

1次アクティブ・サーバ、代替サーバ、およびアービタ・サーバが定義されているトライアッド構成に前記サーバがなっていることを前記情報が示し、その役割が、前記1次アクティブ・サーバ、前記代替サーバ、前記アービタ・サーバ、または前記1次アクティブ・サーバ、前記代替サーバ、あるいは前記アービタ・サーバのいずれでもないサーバであることを前記サーバが判断する、請求項1記載の方法。

【請求項7】

前記判断が、前記サーバに維持された階層1構成情報ブロック内の情報に基づいて行われる、請求項6記載の方法。

30

【請求項8】

前記サーバにより、ある条件の発生に応答して、前記協定タイミング・ネットワーク内の判断されたその役割に基づいて、アクションを実行するステップをさらに含む、請求項1記載の方法。

【請求項9】

前記条件が、デュアル構成またはトライアッド構成内の1次アクティブ・サーバの障害であり、前記サーバが前記故障中の1次アクティブ・サーバ用の1次アクティブ・サーバとして引き継ぐ、請求項8記載の方法。

【請求項10】

前記サーバが、時刻同期を維持するように構成された協定タイミング・ネットワークを介して1次アクティブ・サーバとネットワーク通信しており、前記サーバが前記協定タイミング・ネットワークによりサーバ・タイム・プロトコル・メッセージを受信する、請求項1記載の方法。

40

【請求項11】

請求項1ないし請求項10のいずれかに記載の方法のすべてのステップを実行するためには適合された手段を含むシステム。

【請求項12】

コンピュータ・システム上で実行されるときに、請求項1ないし請求項10のいずれかに記載の方法のすべてのステップを実行するための命令を含むコンピュータ・プログラム。

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、一般に、処理装置のネットワーク内およびそれを越える時刻同期に関し、特に、協定タイミング・ネットワーク内のサーバが同じルート1次基準時間に同期できるようにする、協定タイミング・ネットワーク用の単一アクティブ1次タイム・サーバを有する階層1構成を定義することに関する。

## 【背景技術】

## 【0002】

パフォーマンスおよびデータ保全性のために、ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社によって提供されるシステムplex (Sysplex) など、共用データにアクセスするコンピューティング・システムは、システム間の最良な場合の通信時間より良好な正確度で時刻 (TOD : time of day) 機構の同期を維持できなければならない。現在、一例では、同期要件を満たすために、IBMシステムplex・タイム (商標) などのタイムが使用される。このタイムは、高価な専用タイミング・リンクと個別の外部ボックスを必要とする。IBM (商標) およびIBMシステムplex・タイム (商標) は、ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社の登録商標である。

10

## 【0003】

ネットワーク・タイミング・プロトコル (NTP : Network Timing Protocol) などのネットワークは、時刻同期を提供するが、ハイエンド・システムの正確度要件を満たさない。NTPでは、すべてのサーバが同じ基準時間に同期することを保証するために、マイクロ秒レベルの正確度を提供する外部タイム・ソースに各サーバがアクセスできる必要がある。これは、このレベルの正確度を提供する外部タイム・サーバに接続する機能を備えていないシステムにとっては問題である。さらに、各システム上のGPSレシーバまたは同様の付属装置の要件は、メンテナンス、セキュリティ、および信頼性の理由から実行不可能なものと見なされる可能性がある。

20

## 【先行技術文献】

## 【特許文献】

## 【0004】

30

【特許文献1】米国特許出願第11/468352号

【特許文献2】米国特許出願第60/887512号

【特許文献3】米国特許出願第60/887544号

【特許文献4】米国特許出願第60/887576号

【特許文献5】米国特許出願第60/887584号

【特許文献6】米国特許出願第60/887586号

【特許文献7】米国特許出願第11/460025号

【特許文献8】米国特許出願第11/223886号

【特許文献9】米国特許出願第11/532168号

【特許文献10】米国特許出願第11/468501号

40

【特許文献11】米国特許出願第11/223878号

【特許文献12】米国特許出願第11/223876号

【特許文献13】米国特許出願第11/223577号

【特許文献14】米国特許出願第11/223642号

## 【非特許文献】

## 【0005】

【非特許文献1】2005年9月発行のIBM資料No. SA22-7832-04「z/Architecture Principles of Operation」

## 【発明の概要】

## 【発明が解決しようとする課題】

50

**【0006】**

上記に基づき、時刻同期の提供を促進する機能が必要である。一例では、タイミング・ネットワーク内のサーバが同じ1次基準時間を使用してそれぞれのクロックを同期させることを保証する、タイミング・ネットワーク内の單一アクティブ階層1サーバを指定する機能が必要である。

**【課題を解決するための手段】****【0007】**

タイミング・ネットワーク内の同期を促進するためのコンピュータ可読プログラム・コード・ロジックを有する少なくとも1つのコンピュータ使用可能媒体を含む装置(*article of manufacture*)の提供により、従来技術の短所は克服され、追加の利点が提供される。このコンピュータ可読プログラム・コード・ロジックは、実行されると、たとえば、タイミング・ネットワーク用の構成を定義するステップであって、その構成がタイミング・ネットワークの1つまたは複数の他のサーバ用のクロック・ソースとして使用される單一サーバを含むステップと、タイミング・ネットワークのサーバにその構成を提供して、1つまたは複数の他のサーバが單一サーバからタイミング情報を入手できるようにするステップとを実行する。

10

**【0008】**

本発明の1つまたは複数の態様に関連する方法およびシステムも本明細書に記載され、請求されている。

**【0009】**

20

追加の特徴および利点は本発明の技法により実現される。本発明のその他の諸実施形態および諸態様は、本明細書に詳細に記載され、請求された発明の一部と見なされる。

**【0010】**

本発明の1つまたは複数の態様は、本明細書の終わりの特許請求の範囲において、詳細に指摘され、例として明瞭に請求される。本発明の上記その他の目的、特徴、および利点は、添付図面に併せて取られた以下の詳細な説明から明らかである。

**【図面の簡単な説明】****【0011】**

【図1】本発明の1つまたは複数の態様を取り入れるための混合協定タイミング・ネットワーク(*mixed coordinated timing network*)の一例を示す図である。

30

【図2】本発明の1つまたは複数の態様を取り入れるためのS T P専用ネットワークの一例を示す図である。

【図3】本発明の一態様による階層(stratum)1構成情報ブロックの一例を示す図である。

【図4】本発明の一態様による新規階層1構成情報ブロックの一例を示す図である。

【図5】本発明の一態様によるヌル構成の図表現である。

【図6】本発明の一態様による单一サーバ階層1構成の図表現である。

【図7】本発明の一態様による单一サーバ階層1構成の作成に関連するロジックの一実施形態を示す図である。

【図8】本発明の一態様によるデュアル・サーバ階層1構成の図表現である。

40

【図9】本発明の一態様によるデュアル・サーバ構成の作成に関連するロジックの一実施形態を示す図である。

【図10】本発明の一態様によるトライアッド(triad)構成の図表現である。

【図11】本発明の一態様によるトライアッド構成の作成に関連するロジックの一実施形態を示す図である。

【図12】本発明の一態様による階層1構成変更要求ブロックの一実施形態を示す図である。

【図13】本発明の一態様による階層1構成変更応答ブロックの一実施形態を示す図である。

【図14】本発明の一態様による構成変更通知要求ブロックの一実施形態を示す図である

50

。

【図15】本発明の一態様による構成変更通知応答ブロックの一実施形態を示す図である。

。

【図16】本発明の一態様によるCTN-ID変更要求ブロックの一実施形態を示す図である。

【図17】本発明の一態様によるCTN-ID変更応答ブロックの一実施形態を示す図である。

【図18】本発明の一態様によるCTNパラメータ読み取り要求ブロックの一実施形態を示す図である。

【図19】本発明の一態様によるCTNパラメータ読み取り応答ブロックの一実施形態を示す図である。 10

【図20】本発明の一態様による故障中のアクティブ階層1サーバのリカバリの実行に関連するロジックの一実施形態を示す図である。

【図21】本発明の一態様によるデュアル・サーバ階層1構成リカバリ手順に関連するロジックの一実施形態を示す図である。

【図22】本発明の一態様によるトライアッド階層1構成リカバリ手順に関連するロジックの一実施形態を示す図である。

【図23】本発明の一態様によるコンソール支援リカバリに関連するロジックの一実施形態を示す図である。 20

【図24】本発明の一態様によるコンソール支援リカバリ結果構成の図表現である。

【図25】本発明の一態様によるトライアッド・リカバリの実行に関連するロジックの一実施形態を示す図である。

【図26】本発明の一態様によるアクティブ階層1チェック停止後の階層1引き継ぎ(takeover)の図表現である。

【図27】本発明の一態様によるアクティブ階層1へのリンクの喪失後の階層1サーバ引き継ぎの図表現である。

【図28】本発明の一態様による単一リンクの喪失・階層1引き継ぎなしの図表現である。

【図29】本発明の一態様によるアクティブ階層1引き継ぎの実行に関連するロジックの一実施形態を示す図である。 30

【図30】本発明の一態様によるアクティブ階層1引き渡し(surrender)の実行に関連するロジックの一実施形態を示す図である。

【図31】本発明の一態様によるサーバ状態要求要求ブロックの一実施形態を示す図である。

【図32】本発明の一態様によるサーバ状態要求応答ブロックの一実施形態を示す図である。

【図33】本発明の一態様によるサーバ状態書き込み要求ブロックの一実施形態を示す図である。

【図34】本発明の一態様によるサーバ状態書き込み応答ブロックの一実施形態を示す図である。 40

【図35】本発明の1つまたは複数の態様を取り入れ、使用するための処理環境の一実施形態を示す図である。

【図36】本発明の一態様による図35のメモリの追加詳細を示す図である。

【図37】本発明の1つまたは複数の態様を取り入れるためのコンピュータ・プログラム(computer program product)の一例を示す図である。

【発明を実施するための形態】

【0012】

本発明の一態様により、タイミング・ネットワーク内のサーバが同じ1次基準時間を使用してそれぞれの時刻機構を同期させることを保証する際に使用可能な單一アクティブ階層1サーバを含むタイミング・ネットワーク用の階層1構成を定義するための機能が提供 50

される。したがって、ネットワーク内のサーバは同じルート1次基準時間に同期しており、同期正確度は階層1サーバにおける外部タイム・ソースの品質または外部タイム・ソースの存在に依存しない。

#### 【0013】

本発明の他の一様では、サーバの保全性を保証するために使用される他の情報とともにネットワークのサーバに提供される階層1構成情報ブロック内に階層1構成が維持される。ネットワーク内の各サーバは階層1サーバのアイデンティティを認識しているので、他の階層1サーバに同期していることを示すサーバはそのネットワークに加入することができない。

#### 【0014】

本発明のさらに他の一様では、ネットワークのアクティブ階層1サーバの障害を処理するためのリカバリ手順が提供される。

#### 【0015】

階層1サーバを含むように様々なネットワークを構成することができるが、このようなネットワークの1つは協定タイミング・ネットワーク（CTN：Coordinated Timing Network）である。協定タイミング・ネットワークでは、複数の別個のコンピューティング・システムが時刻同期を維持して、協定タイミング・ネットワークを形成する。協定タイミング・ネットワーク内のシステムは、サーバ・タイム・プロトコル（STP：Server Time Protocol）というメッセージ・ベース・プロトコルを使用して、既存の高速データ・リンクによりシステム間で計時情報を受け渡す。これにより、各システムの時刻（TOD）機構を現代のハイエンド・コンピューティング・システムに必要な正確度で同期させることができる。このプロトコルはコンピューティング・システム内の技術を使用するので、技術が向上するにつれて同期正確度が高くなる。STP機能を提供するコンピューティング・システムは本明細書ではタイム・サーバまたはサーバという。

#### 【0016】

1次タイム・サーバとしてCTNにおいて定義されたサーバは、CTN用の1次基準時間を提供する。CTN内の他のサーバからの情報に基づいてCST（CTN用の時刻（TOD）機構の推定値）を決定するCTN内のサーバは2次タイム・サーバという。1次タイム・サーバは外部タイム・ソースからその時刻を入手することができ、その外部タイム・ソースはCTN内の時刻機構を定義済み時間標準に同期させるための手段を提供する。

#### 【0017】

同期状態にあるCTN内のサーバには、それと1次タイム・サーバとの間のサーバの数を指定する、階層レベルという値が割り当てられる。1次タイム・サーバは1という階層レベルで動作し、2次タイム・サーバは2以上の階層レベルで動作し、これは階層1までのタイミング経路内のサーバの数が増加するにつれて上昇する。一般に、階層レベルが上昇するにつれて、計時情報の品質が低下する。同期していないサーバには0という階層レベルが割り当てられる。

#### 【0018】

STP機能は、STPメッセージを送信し、受信し、処理するために必要な手順を提供する。STPメッセージは、サーバ間の1つまたは複数の物理データ・リンクにより送信される。2つのサーバ間に確立されたデータ・リンクはSTP経路という。STP機能は、STP経路を確立し維持するための諸機能を提供する。

#### 【0019】

STPメッセージは、メッセージ・コマンドとメッセージ応答とを含む。時間パラメータ交換（XTP：exchange time parameters）メッセージとSTP制御（STC：STP control）メッセージという2つのタイプのSTPメッセージがサポートされる。XTPメッセージは、CTN用のCSTを決定するために使用される計時情報を交換するために使用される。STP制御メッセージは、CTN内のサーバによって必要とされる様々なCTNパラメータを設定し変更するために使用される。

#### 【0020】

10

20

30

40

50

CTNは、たとえば、2つのタイプの構成のうちの1つとして、すなわち、混合CTN構成またはSTP専用CTN構成のいずれかとして動作することができる。混合CTN構成では、サーバは、STPネットワークと外部時間基準(ETR:External Time Reference)ネットワークの両方の一部になるように構成される。混合CTN構成では、CTN内のサーバは同じ非ヌルETRネットワークIDを備えて構成され、タイマ(たとえば、IBMシステム・タイマ(商標))はCTN用の1次時間基準を提供する。CTN内の少なくとも1つのサーバは、CTN内で同期を行えるようになる前にシステム・タイマによって提供されたタイミング信号に合わせて進む予定である。システム・タイマに合わせて進まないサーバは2次タイム・サーバであり、STP信号を交換することによって同期を達成する。

10

#### 【0021】

一例として、システム・タイマのタイミング信号に合わせて進む各サーバはタイム・プロトコル・パラメータを受信し、たとえば、CTNパラメータ更新手順を使用して、CTN内の2次タイム・サーバにその情報を伝搬する。この手順の一例は、Carls on他により2006年8月30日に出願され、「Coordinated Timing Network Configuration Parameter Update Procedure」という名称の米国特許出願第11/468352号に記載されている。

#### 【0022】

図1に関連して、混合CTN構成100の一例について説明する。混合CTN構成100は、たとえば、ローカル・エリア・ネットワーク(104)に結合されたサーバA(102)と、ローカル・エリア・ネットワーク(104)に結合されたサーバB(106)と、ローカル・エリア・ネットワーク(110)に結合されたサーバC(108)とを含む。各サーバは、たとえば、インターナショナル・ビジネス・マシーンズ社によって提供されるz/Architecture(商標)に基づく中央演算処理複合システムである。z/Architecture(商標)は米国ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社の登録商標である。z/Architecture(商標)の一実施形態は、2005年9月発行のIBM資料No. SA22-7832-04「z/Architecture Principles of Operation」に記載されている。

20

#### 【0023】

各ローカル・エリア・ネットワークは、ネットワーク内で時刻同期を提供する際に使用されるコンソール120に結合される。さらに、ローカル・エリア・ネットワーク104とローカル・エリア・ネットワーク110は、広域ネットワーク112を介して相互に結合される。

30

#### 【0024】

サーバAおよびBは外部時間基準ネットワーク114に結合され、サーバBおよびCはSTPネットワーク116の一部になるように構成される。サーバBは階層1レベルであり、サーバCは階層2レベルである。STPリンク118は、サーバBのSTP機能をサーバCのSTP機能に結合するために使用される。

#### 【0025】

STP専用CTNでは、CTN内のサーバはSTPネットワークの一部になるように構成され、ETRネットワークの一部になるように構成されるものは何もない。図2に関連して、STP専用ネットワーク150の一例について説明する。この例では、サーバA(152)およびサーバB(154)はLAN(156)に結合され、サーバC(158)はLAN(160)に結合される。サーバのそれぞれはSTP機能162を含み、各機能は1つまたは複数のSTPリンク164を介して相互に結合される。

40

#### 【0026】

さらに、LAN156はコンソール170に結合され、LAN160はコンソール172に結合される。コンソール170は、テレフォン・タイム・サーバ(たとえば、ACTS:NIST自動コンピュータ・タイム・サービス)へのダイヤルアウトなどの外部タイム・ソース(ETS:external time source)174にさらに結合される。このネットワ

50

ークには、ETRネットワークはまったく存在しない。サーバBは1という階層レベルを有し、サーバAおよびCは2という階層レベルを有する。

【0027】

STP専用ネットワークなどのネットワーク内でアクティブ階層1サーバとして動作する予定のサーバは、そのネットワーク用に定義された階層1構成の一部として指定される。階層1構成は、そのネットワークの各サーバで維持され、たとえば、そのネットワーク用に定義された構成のタイプを含む、そのネットワークの構成に関する情報を提供する。ネットワークは、たとえば、以下のものを含む様々なタイプの1つとして構成することができる。

a) ヌル構成 - ヌル構成では、階層1サーバは識別されない。そのサーバは、非ヌル階層1構成を有するサーバに接続するまで、非同期のままになる。階層レベル0にあるサーバの階層1構成は、たとえば、任意の他のサーバに接続されないときにヌル構成に等しくなり、後述する階層1構成情報ブロック内の单一CEC-CTNインジケータは0になる。

b) 単一サーバ定義 - 単一サーバ定義では、階層1構成は、CTN用のアクティブ階層1サーバとして動作する单一1次階層1サーバを定義する。1次階層1サーバが失われると、新規階層1構成がそのコンソールから指定されるまで、CTN内の同期時刻が失われることになる。

单一サーバ階層1構成は、指定の階層1サーバがCTN内で唯一のサーバ（たとえば、コンピューティング電子複合システム（CEC）またはCPC）であり、他のサーバまたは追加のサーバはいずれもCTNの一部にならないという表示も含むことができる。この場合のCTNは、单一CEC-CTNという。それが单一CEC-CTNであることを階層1構成が示す場合、その構成は、パワーオン・リセット後のサーバの初期化状態のために有効な階層1構成である。それが单一CEC-CTNであることを单一サーバ階層1構成が指定しない場合、その構成は、パワーオン・リセット後のサーバの初期化状態のために有効な階層1構成ではなく、階層1構成はヌル構成に設定される。

c) デュアル・サーバ構成 - デュアル・サーバ構成では、この構成は1次階層1サーバと代替階層1サーバとを含む。デュアル・サーバ構成を使用すると、CTN用のアクティブ階層1サーバの役割を代替サーバが引き継ぐためのメカニズムを提供する。代替階層1サーバは、CTN内の他のサーバに対して1次階層1サーバと同じ接続性を備えて構成されると、CTNの同期機能に対する妨害なしに、アクティブ階層1として引き継ぐことができる。CTN用のアクティブ階層1サーバとして動作している階層1サーバ（1次または代替）は、後述するように、CTN内の各サーバに維持される階層1構成情報ブロック内に示される。アクティブ階層1サーバではない階層1サーバは、非アクティブ階層1サーバとして定義される。

デュアル・サーバ構成内の非アクティブ階層1サーバは、アクティブ階層1障害を検出すると、アクティブ階層1サーバの役割を引き継ぐ。以下のうちの1つが発生すると、デュアル・サーバ構成内でアクティブ階層1障害が検出される。

コンソール支援リカバリ手順が実行され、その手順により、アクティブ階層1障害が発生したことが示されるか、または

階層1システム・チェック信号が認識される。

デュアル・サーバ構成内の非アクティブ階層1サーバは、その非アクティブ階層1サーバがアクティブ階層1サーバへの接続を失うと、コンソール支援リカバリを実行する。

d) トライアッド構成 - トライアッド構成では、階層1構成は、1次階層1サーバと、代替階層1サーバと、アービタ・サーバ(arbiter server)とを含む。トライアッド構成の定義は、デュアル・サーバ構成用に定義されたように、CTN用のアクティブ階層1サーバの役割を代替サーバが引き継ぐためのメカニズムを提供する。さらに、1次階層1サーバと代替階層1サーバの両方への接続性を備えて構成されると、アービタ・サーバの定義は、非アクティブ階層1サーバとアービタ・サーバが相互に通信して、アクティブ階層1サーバ障害が発生していることと、非アクティブ階層1サーバがアクティブ階層1

10

20

30

40

50

サーバとして引き継がなければならないことを判断するためのメカニズムを提供する。

トライアッド構成内の非アクティブ階層1サーバは、アクティブ階層1障害を認識すると、アクティブ階層1の役割を引き継ぐ。非アクティブ階層1サーバは、以下のいずれかが発生すると、アクティブ階層1障害を認識する。

トライアッド・リカバリ手順が実行され、その手順により、アクティブ階層1障害が発生したことが示される。

コンソール支援リカバリ手順が実行され、その手順により、アクティブ階層1障害が発生したことが示される。

非アクティブ階層1サーバは、以下の条件が発生すると、トライアッド・リカバリ手順を実行する。

非アクティブ階層1サーバがアクティブ階層1サーバへの接続を失い、アービタへの接続が可能である。

非アクティブ階層1サーバがアクティブ階層1通信タイムアウトを認識し、アービタへの接続が可能である。

非アクティブ階層1サーバは、その非アクティブ階層1サーバがアクティブ階層1サーバへの接続を失い、アービタに接続されていないときに、コンソール支援リカバリを実行する。

非アクティブ階層1サーバは、CTN用のアクティブ階層1サーバの役割を引き継ぐためにアクティブ階層1引き継ぎ手順を実行する。

トライアッド構成内のアクティブ階層1サーバは、非アクティブ階層1サーバとアービタ・サーバの両方への接続を失ったことを検出すると、アクティブ階層1サーバの役割を断念する。アクティブ階層1サーバは、CTN用のアクティブ階層1サーバの役割を断念するためにアクティブ階層1引き渡し手順を実行する。

#### 【0028】

一例では、階層1構成情報は、階層1構成情報ブロック (SCIB : stratum-1 configuration information block) という制御ブロック内に維持され、そのブロックはネットワークの各サーバ上に保管されるかまたは各サーバにとってアクセス可能なものである。SCIBは、ネットワーク用の階層1構成を識別するために使用される。

#### 【0029】

図3に関連して、階層1構成情報ブロック200の一実施形態について説明する。階層1構成情報ブロック200は、たとえば、以下のフィールドを含む。

a) 1次階層1ノード記述子202 : このフィールドは、後述する構成情報ブロックの構成タイプ・フィールドに単一サーバ定義、デュアル・サーバ定義、またはトライアッド定義が指定されたときに有効であり、有効である場合、1次階層1ノード記述子のノード記述子を含む。

b) 代替階層1ノード記述子204 : このフィールドは、構成タイプ・フィールドにデュアル・サーバ定義またはトライアッド定義が指定されたときに有効であり、有効である場合、代替階層1サーバ・ノード記述子のノード記述子を含む。

c) アービタ・ノード記述子206 : このフィールドは、構成タイプ・フィールドにトライアッド定義が指定されたときに有効であり、アービタ・ノード記述子のノード記述子を含む。

d) 階層1構成タイムスタンプ208 : このフィールドは、このブロック内の階層1構成情報がサーバで最新 (current) のものになった時間を示すタイムスタンプを含む。

e) 構成タイプ (Type) 210 : このフィールドは、以下に定義されるように、階層1構成のタイプを指定する。

ヌル定義 : いずれのノード記述子も有効ではない。

単一サーバ定義 : 1次階層1ノード記述子のみが有効である。

デュアル・サーバ定義 : 1次階層1ノード記述子および代替階層1ノード記述子が有効である。

トライアッド定義 : 1次階層1ノード記述子、代替階層1ノード記述子、および

10

20

30

40

50

アービタ・ノード記述子が有効である。

f ) アクティブ階層 1 ( A ) 212 : このフィールドは、デュアル・サーバ定義またはトライアッド定義が指定されたときに有効であり、1次階層 1 サーバまたは代替階層 1 サーバがアクティブ階層 1 サーバであるかどうかを示す。

g ) 単一 C E C C T N ( X ) 214 : このフィールドは、構成タイプが单一サーバ定義を指定したときに有効であり、一例として、1である場合、C T N が单一 C E C C T N であることを示す。このフィールドがたとえば 0 である場合、C T N は单一 C E C C T N ではない。

h ) リカバリ構成 ( R ) 216 : このフィールドは、このブロックによって記述された階層 1 構成が階層 1 リカバリ・アクションまたはコンソール・コマンドの結果であるかどうかを示す。 10

i ) 階層 1 最大短期スキュー・レート ( skew rate ) 変更 218 : このフィールドは、任意の指定の期間 ( たとえば、60秒 ) に発生する可能性のある、階層 1 サーバの物理的発振器のスキュー・レートの最大可能変更を指定する値を含む。このフィールドは、ベース・ステアリング・レート ( base-steering rate ) と同じフォーマットの値を形成するために使用される。

動的発振器切り替えを実行できないマシンでは、この値は、指定の期間 ( たとえば、60秒 ) に発生する可能性のある最大階層 1 発振器周波数ドリフトに等しい。動的発振器切り替えを実行できるマシンでは、この値は、その発振器について指定された最大範囲のスキュー許容度に設定される。たとえば、発振器スキュー許容度が公称周波数に対して  $\pm 2 \text{ ppm}$  に指定された動的発振器切り替えをサポートするマシンでは、この値は  $4 \text{ ppm}$  という等価量に設定される。システム T O D 機構を駆動するために使用される物理的発振器がある発振器から他の発振器に切り替えられると、動的発振器切り替えが行われる。 20

### 【 0030 】

上記の制御ブロックに加えて、新規階層 1 構成情報ブロック ( N S C I B : new stratum-1 configuration information block ) という他の制御ブロックを使用して、C T N 用の新規階層 1 構成を指定することができる。さらに、これは、階層 1 構成変更と同時に行われる予定の C T N I D の更新を指定するために使用することができる。

### 【 0031 】

一例では、あるサーバの N S C I B は、そのサーバが S T P 専用 C T N 構成の一部になるように構成されたときまたは N S C I B 内の S T P マイグレーション・ビットが 1 に等しい場合に意味を持つ。 30

### 【 0032 】

図 4 に関連して、新規階層 1 構成情報ブロック 300 の一実施形態について説明する。新規階層 1 構成情報ブロック 300 は、たとえば、以下のフィールドを含む。

a ) 1次階層 1 ノード記述子 302 : このフィールドは、構成タイプ・フィールドに单一サーバ定義、デュアル・サーバ定義、またはトライアッド定義が指定されたときに有効であり、新規 1次階層 1 ノード記述子のノード記述子を含む。

b ) 代替階層 1 ノード記述子 304 : このフィールドは、構成タイプ・フィールドにデュアル・サーバ定義またはトライアッド定義が指定されたときに有効であり、新規代替階層 1 ノード記述子のノード記述子を含む。 40

c ) アービタ・ノード記述子 306 : このフィールドは、構成タイプ・フィールドにトライアッド定義が指定されたときに有効であり、新規アービタ・ノード記述子のノード記述子を含む。

d ) 階層 1 構成更新時間 308 : サーバが S T P 専用 C T N の一部になるように構成されたときに、このフィールドは、このブロック内の値が C T N について最新のものになる時期を示すタイムスタンプを含む。

e ) 構成タイプ ( C T y p e ) 310 : このフィールドは、以下に定義されるように、階層 1 構成のタイプを指定する。

ヌル定義 : いずれのノード記述子も有効ではない。

単一サーバ定義： 1次階層1ノード記述子のみが有効である。

デュアル・サーバ定義： 1次階層1ノード記述子および代替階層1ノード記述子が有効である。

トライアッド定義： 1次階層1ノード記述子、代替階層1ノード記述子、およびアービタ・ノード記述子が有効である。

f) アクティブ階層1(A)312: このフィールドは、デュアル・サーバ定義またはトライアッド定義が指定されたときに有効であり、1次階層1サーバまたは代替階層1サーバがアクティブ階層1サーバであるかどうかを示す。

g) CTN\_ID変更(C)314: サーバがSTP専用CTNの一部になるように構成されたときに、このフィールドは、CTN\_ID変更が要求されているかどうか、ならびにCTN\_IDが有効であるかどうかを示す。この変更は階層1構成更新時間に行われる。

CTN\_IDは、CTNを識別するために使用される値である。CTN\_IDは、たとえば、STPネットワークIDとETRネットワーク番号とを含む。STPネットワークIDは、このサーバ用に構成されたSTPネットワークがある場合にそのネットワークを識別するものである。ETRネットワーク番号は、このサーバ用に構成されたETRネットワークがある場合にそのネットワークを識別するものである。

h) STPマイグレーション構成(S)316: このフィールドは、たとえば、このサーバがSTP専用CTNとして構成されない場合に意味を持つ。このフィールドは、このサーバについてSTP専用マイグレーション階層1構成が定義されているかどうかを示す。このフィールドがたとえば1である場合、NSCIBは、このサーバのCTN\_IDがSTP専用構成に変更されたときに最新のものになる予定の階層1構成を含む。このフィールドがたとえば0である場合、このサーバについてSTP専用マイグレーション階層1構成は定義されない。

i) 単一CEC\_CTN(X)318: このフィールドは、構成タイプが単一サーバ定義を指定したときに有効であり、一例として、1である場合、CTNが単一CEC\_CTNであることを示す。このフィールドがたとえば0である場合、CTNは単一CEC\_CTNではない。

j) リカバリ構成(R)320: このフィールドは、このブロックによって記述された階層1構成が階層1リカバリ・アクションまたはコンソール・コマンドの結果であるかどうかを示す。

k) 階層1最大短期スキューレート変更322: このフィールドは、任意の指定の期間(たとえば、60秒)に発生する可能性のある、新規階層1サーバの物理的発振器のスキューレートの最大可能変更を指定する値を含む。

l) 新規CTN\_ID320: このフィールドは、CTN\_ID変更インジケータがCTN\_ID変更要求および有効なCTN\_IDを指定したときに有効である。このフィールドは新規CTN\_IDを指定するものである。

#### 【0033】

CTN\_IDの更新を指定するために新規階層1構成情報ブロックが使用されない場合、このブロックは、一例として、CTN\_ID変更ビットまたは新規CTN\_IDを含まない可能性がある。

#### 【0034】

階層1構成は、マシンの初期パワーオン時にサーバにおいてヌル構成に初期化される。ヌル構成の一例は図5に示されている。描寫されているように、ヌル構成400は複数のサーバ402を含み、これらのサーバは1つまたは複数のSTPリンク404を介して相互に結合される。この例では、各サーバの階層レベルは0に等しく、したがって、アクティブ階層1サーバはまったく存在しない。しかし、CTN\_IDは特定の値に設定され、その値は、この例では、後述するようにローカルCTN\_ID変更コンソール・コマンドを使用して設定される。

#### 【0035】

10

20

30

40

50

ヌル階層 1 構成を有する C T N は、階層 1 が指定されておらず、その C T N 用の基準時間がまったくないので、同期させることができない。C T N 用の階層 1 構成は、後述する階層 1 構成変更コマンドなどのコマンドを介して非ヌル構成からヌル構成に設定を戻すことができる。

【 0 0 3 6 】

ヌル構成から、单一サーバ階層 1 を作成することができる。单一サーバ階層 1 構成は、单一階層 1 サーバを指定するものであって、代替階層 1 サーバまたはアービタ・サーバはまったく指定しない。单一サーバ階層 1 構成の一例は図 6 に描写されている。図示の通り、单一サーバ階層 1 構成 5 0 0 は、1 つまたは複数の S T P リンク 5 0 4 を介して相互に結合された複数のサーバ 5 0 2 を含み、そのサーバのうちの 1 つ、たとえば、サーバ A はアクティブ 1 次階層 1 サーバとして選択される。したがって、サーバ A は 1 次 S 1 という表示を有し、その階層レベルは 1 に等しい。他のサーバは、2 または 3 など、1 より大きい階層値を有することに留意されたい。

【 0 0 3 7 】

図 7 に関連して、单一サーバ階層 1 構成の作成に関連するロジックの一実施形態について説明する。最初に、ステップ 6 0 0 で、コンソールは、階層 1 構成変更コマンドなどのコマンドをサーバ A に発行し、1 次階層 1 サーバとしてサーバ A を有する单一サーバ階層 1 構成を指定する。

【 0 0 3 8 】

その後、ステップ 6 0 2 で、サーバ A は、新規階層 1 構成が最新のものであることを示す。一例では、サーバ A は、構成変更通知コマンドをコンソールに発行して、この表示を提供する。

【 0 0 3 9 】

その次に、ステップ 6 0 4 で、ネットワークの他のサーバは、ロック・ソースの可用性を認識し、C T N に関する現行階層 1 構成および時間制御パラメータを入手する。一例では、これは、サーバが C T N パラメータ読み取り S T P 制御メッセージをロック・ソースに発行して、C T N に関する現行階層 1 構成および時間制御パラメータを入手することによって実行される。

【 0 0 4 0 】

さらに、ステップ 6 0 6 で、サーバは、構成変更マシン・チェックを発行し、一例では、新規階層 1 構成を示す。また、ステップ 6 0 8 で、サーバは、たとえば、タイミング・アラート外部割り込みを発行することにより、タイミング状況変更（使用可能ロック・ソース、同期）も示す。これで、ヌル構成からの单一サーバ階層 1 構成の作成に関連する処理を終了する。

【 0 0 4 1 】

上記に加えて、既存の非ヌル構成からデュアル・サーバ構成を作成することができる。デュアル・サーバ階層 1 構成は、1 次階層 1 サーバおよび代替階層 1 サーバを指定するものであり、そのうちの一方が C T N 用のアクティブ階層 1 サーバとして動作することができる。階層 1 サーバとして動作する予定のサーバは、階層 1 構成が作成されたときにコンソールによって指定され、任意の時期にコンソールによって変更することができる。

【 0 0 4 2 】

デュアル・サーバ階層 1 構成の一例は図 8 に描写されている。図示の通り、サーバ A (7 0 0) は、1 次として示され、1 に等しい階層レベルを有する。サーバ C (7 0 2) は、代替として示され、2 に等しい階層レベルを有する。サーバ B (7 0 4) およびサーバ D (7 0 6) は 2 に等しい階層レベルを有し、サーバ E (7 0 8) は 3 に等しい階層レベルを有する。

【 0 0 4 3 】

図 9 に関連して、デュアル・サーバ構成の作成に関連するロジックの一実施形態について説明する。最初に、ステップ 8 0 0 で、1 次およびアクティブとしてサーバ A を有し、代替としてサーバ C を有する、デュアル・サーバ階層構成が指定される。一例では、これ

10

20

30

40

50

は、コンソールが階層 1 構成変更コマンドをサーバ A に発行することによって実行される。

【 0 0 4 4 】

その後、照会 8 0 2 で、階層 1 構成変更コマンド内の強制変更ビットが 0 であるかどうかについて判断が行われる。それが 0 である場合、ステップ 8 0 4 で、サーバ A は、それがサーバ C に直接接続されていることを検証する。一例では、サーバ C とメッセージを交換することによって、この接続が検証される。照会 8 0 6 で、サーバ A がサーバ C に直接接続されていることが検証によって示された場合、ステップ 8 0 8 で、サーバ A が C T N 内の階層 1 構成更新をスケジュールし、ステップ 8 0 9 で、更新が行われる。さらに、ステップ 8 1 0 で、新規階層 1 構成が最新のものになったときに、サーバ A が一例として構成変更通知コマンドをコンソールに発行する。また、ステップ 8 1 2 で、ネットワークのサーバは、新規階層 1 構成が最新のものになったことに応答して、構成変更マシン・チェックも発行する。

【 0 0 4 5 】

照会 8 0 2 に戻り、強制変更ビットが 1 に設定された場合、処理はステップ 8 0 8 を続行し、そこでサーバ A は C T N 内の階層 1 構成更新をスケジュールする。

【 0 0 4 6 】

もう一度、照会 8 0 6 に戻り、サーバ A がサーバ C に接続されていない場合、ステップ 8 1 4 でコマンドが拒否され、処理が完了する。

【 0 0 4 7 】

既存の非ヌル構成からデュアル構成を作成することに加えて、既存の非ヌル構成からトライアッド構成を作成することもできる。トライアッド階層 1 構成は、1 次階層 1 サーバ、代替階層 1 サーバ、およびアービタ・サーバを指定する。指定の 1 次または代替のいずれか一方が C T N 用のアクティブ階層 1 サーバとして動作することができる。階層 1 構成が作成されたときにコンソールによって指定された階層 1 サーバとして動作する予定のサーバは、任意の時点でコンソールによって変更することができる。アービタ・サーバは、アクティブ階層 1 サーバがもはや C T N の一部ではない時期の決定を支援するために使用される。

【 0 0 4 8 】

トライアッド構成の一例は図 1 0 に描写されている。図示の通り、サーバ A ( 9 0 0 ) は 1 次階層 1 サーバであり、1 という階層レベルを有する。さらに、サーバ C ( 9 0 2 ) は代替階層 1 サーバであり、2 という階層レベルを有し、サーバ B ( 9 0 4 ) はアービタであり、これも 2 という階層レベルを有する。さらに、サーバ D ( 9 0 6 ) は 2 という階層レベルを有し、サーバ E ( 9 0 8 ) は 3 という階層レベルを有する。

【 0 0 4 9 】

図 1 1 に関連して、トライアッド構成の作成に関連するロジックの一実施形態について説明する。最初に、ステップ 1 0 0 0 で、コンソールは、たとえば、階層 1 構成変更コマンドをサーバ A に発行し、サーバ A がアクティブ 1 次階層 1 サーバであり、サーバ C が代替階層 1 サーバであり、サーバ B がアービタであるトライアッド構成を指定する。

【 0 0 5 0 】

その後、照会 1 0 0 2 で、階層 1 構成変更コマンド内の強制変更ビットが 0 に設定されているかどうかについて判断が行われる。それが 0 に設定された場合、ステップ 1 0 0 4 で、サーバ A は、それがサーバ C および B に直接接続されていることを検証する。照会 1 0 0 6 で、サーバ A がサーバ C および B に接続されている場合、ステップ 1 0 0 8 で、サーバ A が C T N 内の階層 1 構成更新をスケジュールし、ステップ 1 0 0 9 で、更新が行われる。また、ステップ 1 0 1 0 で、サーバ A は、新規階層 1 構成が最新のものになったことに応答して、構成変更通知コマンドをコンソールに発行する。その後、ステップ 1 0 1 2 で、新規階層 1 構成が最新のものになったときに、サーバは構成変更マシン・チェックを発行する。

【 0 0 5 1 】

10

20

30

40

50

照会 1002 に戻り、変更ビットが 1 に等しい場合、接続性検証は実行されず、処理はステップ 1008 を続行し、そこでサーバ A は CTN 内の階層 1 構成更新をスケジュールする。

#### 【0052】

さらに、照会 1006 で、接続性検証が失敗した場合、ステップ 1014 でコマンドが拒否され、処理が完了する。

#### 【0053】

上述したように、非ヌル構成を作成する際に、様々なコマンドが使用される。このようなコマンドの 1 つは階層 1 構成変更コマンドであり、これは CTN 用の階層 1 構成に対する変更を指定するために使用される。このコマンド要求は、階層 1 構成を変更する方法を指定する新規階層 1 構成情報ブロックを含む。

10

#### 【0054】

図 12 に関連して、階層 1 構成変更コマンド要求ブロック 1100 の一実施形態について説明する。一例では、階層 1 構成変更コマンド要求ブロックは以下のフィールドを含む。

- a) 長さフィールド 1102 : このフィールドは、この要求ブロックの長さを示す。
- b) コマンド・コード 1104 : このフィールドは、階層 1 構成変更コマンドを指定する。
- c) コマンド・トランザクション番号 1106 : このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。
- d) CTN ID 1108 : このフィールドは、そのコマンドのターゲットになる CTN の CTN ID を含む。
- e) 強制変更 (F) 1110 : このフィールドは、そのコマンドについて構成検証を実行すべきかどうかを示す。
- f) 新規階層 1 構成情報ブロック 1112 : このフィールドは、その CTN 用の新規階層 1 構成情報ブロックを含む。

20

#### 【0055】

図 13 に関連して、階層 1 構成変更コマンド用の応答ブロックの一実施形態について説明する。階層 1 構成変更応答ブロック 1150 は、たとえば、以下のフィールドを含む。

30

- a) 長さ 1152 : このフィールドは、コマンド応答ブロックの長さを含む。
- b) 応答コード 1154 : このフィールドは、そのコマンド用の応答コードを含む。
- c) コマンド・トランザクション番号 1156 : このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供される値を含む。

#### 【0056】

動作時に、コマンドが CTN 用の新規アクティブ階層 1 サーバを指定すると、NSCI B 内にアクティブ階層 1 サーバとして指定されたサーバにコマンドが発行される。それがそのサーバに発行されない場合、グローバル・コマンド拒否応答コードが返される。

#### 【0057】

その更新が CTN 内で効力を発した後、新規アクティブ階層 1 サーバは非同期的に構成変更通知コマンドを発行する。そのコマンドが混合 CTN 内のサーバに発行されると、NSCI B 内の STP マイグレーション (S) フィールドが 1 に設定される。このフィールドは、CTN 構成が混合 CTN から STP 専用構成に変更されたときに階層 1 構成が効力を発する予定であることを示す。

40

#### 【0058】

コマンド要求ブロックは、そのコマンドについて構成検証を実行すべきかどうかを示すためのフィールドを含む。構成検証なしにそのコマンドを実行する予定であることを示すために、強制変更フィールドは、たとえば、1 に設定される。そのコマンドを受け入れる前に以下の構成検証チェックを実行する予定であることを示すために、強制変更フィールドは、たとえば、0 に設定される。

50

非ヌル構成が指定された場合、受信側サーバが階層 0 サーバではないことを検証する。階層 0 サーバをアクティブ階層 1 サーバに設定するために、強制変更フィールドは 1 に設定されるはずである。

デュアル構成が指定された場合、受信側サーバは以下のことを検証する。

受信側サーバが 1 次階層 1 サーバとして指定された場合、これは、それが代替階層 1 サーバに接続されることを保証する。

受信側サーバが代替階層 1 サーバとして指定された場合、これは、それが 1 次階層 1 サーバに接続されることを保証する。

トライアッド構成が指定された場合、受信側サーバは以下のことを検証する。

受信側サーバが 1 次階層 1 サーバとして指定された場合、これは、それが代替階層 1 サーバおよびアービタ・サーバに接続されることを保証する。 10

受信側サーバが代替階層 1 サーバとして指定された場合、これは、それが 1 次階層 1 サーバおよびアービタ・サーバに接続されることを保証する。

検証チェックが失敗した場合、そのコマンドは受け入れられず、応答コードが報告される。

#### 【 0 0 5 9 】

上述したもう 1 つのコマンドは、構成変更通知 ( C C N : configuration change notification ) コマンドである。構成変更通知コマンドは、 C T N 内の特定の構成変更をコンソールに通知するために使用される。このコマンドは、その構成が C T N 内で効力を発した後にネットワークのサーバによって発行される。このコマンドは、以下の構成パラメータ変更、例として、階層 1 構成変更、階層 1 構成変更および C T N I D 変更、または C T N I D 変更、あるいはこれらの組み合わせについて発行される。 20

#### 【 0 0 6 0 】

構成変更は、階層 1 構成変更コマンドまたはグローバル C T N I D 変更コマンドの結果として、あるいは C T N 用のアクティブ階層 1 サーバの変更を引き起こす C T N リカバリ・アクションの結果として行われる。

#### 【 0 0 6 1 】

図 14 に関連して、構成変更通知コマンド用のコマンド要求ブロック 1 2 0 0 の一例について説明する。一実施形態では、コマンド要求ブロック 1 2 0 0 は以下のものを含む。

- a ) 長さ 1 2 0 2 : このフィールドは、コマンド・ブロックの長さを含む。
- b ) コマンド・コード 1 2 0 4 : このフィールドは、構成変更通知コマンドの指定を含む。
- c ) コマンド・トランザクション番号 1 2 0 6 : このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。

d ) C T N I D 1 2 0 8 : このフィールドは、そのコマンドを送信しているサーバの C T N I D を含む。

e ) 変更コード : このフィールドは、報告されている構成変更を指定するコードを含む。このフィールドは、たとえば、以下のものを示す。

階層 1 構成変更 : 階層 1 構成は、現行階層 1 構成ブロックによって示された通りに変更された。 40

階層 1 構成変更および C T N I D 変更 : 階層 1 構成および C T N I D は、現行階層 1 構成ブロックおよび C T N I D フィールドによって示された通りに変更された。

C T N I D 専用変更 : C T N I D は、 C T N I D フィールドによって示された通りに変更された。

f ) 理由コード : このフィールドは、階層 1 構成変更が報告されている理由を指定するコードを含む。このフィールドは、階層 1 構成変更が変更コードに示されているときに有効であり、例として、以下のうちの 1 つを示す。

階層 1 構成変更 : 階層 1 構成変更は、階層 1 構成変更コマンドの結果である。 50

リカバリ：階層1構成変更は、S T P機能リカバリ・アクションの結果である。

g) 現行階層1構成情報ブロック：このフィールドは、そのサーバ用の現行階層1構成情報ブロックを含む。

#### 【0062】

図15に関連して、構成変更通知コマンド用のコマンド応答ブロック1250の一実施形態について説明する。一例では、コマンド応答ブロック1250は、以下のものを含む。

a) 長さ1252：このフィールドは、コマンド応答ブロック長を指定する。

b) 応答コード1254：このフィールドは、そのコマンド用の応答コードを含む。

c) コマンド・トランザクション番号1256：このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供される値を含む。 10

#### 【0063】

実行できるもう1つのコマンドはC T N I D変更コマンドであり、これは、ローカルまたはグローバルC T N I D変更を実行するために使用される。このコマンドは、C T NのS T PネットワークI Dコンポーネント、E T RネットワークI Dコンポーネント、またはC T N I Dの両方のコンポーネントを変更するために使用することができる。要求ブロック内の有効性インジケータは、C T N I Dのどちらのコンポーネントを変更する予定であるかを示す。要求ブロック内のグローバル・コマンド・パラメータは、そのコマンドがグローバル・コマンドであるかローカル・コマンドであるかを指定する。 20

#### 【0064】

図16に関連して、C T N I D変更コマンド用のコマンド要求ブロック1300の一例について説明する。C T N I D変更要求ブロック1300は、たとえば、以下のものを含む。

a) 長さ1302：このフィールドは、コマンド・ブロック長を含む。

b) コマンド・コード1304：このフィールドは、C T N I D変更コマンドを指定する。

c) コマンド・トランザクション番号1306：このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。

d) C T N I D1307：このフィールドは、有効である場合、そのコマンドのターゲットになるC T NのC T N I Dを含む。C T N I Dは、グローバルC T N更新インジケータがたとえば1であるときに有効であり、受信側サーバの現行C T N I Dと一致する予定であり、そうではない場合、C T N I D不一致エラーが認識される。グローバルC T N更新インジケータがたとえば0である場合、このフィールドは有効ではなく、無視される。 30

e) グローバルC T N更新(G)1308：このフィールドは、コマンドがグローバル更新であるかローカル更新であるかを示す。コマンドがローカル更新である場合、それは受信側サーバでのみ実行される。コマンドがグローバル・コマンドである場合、そのコマンドはアクティブ階層1サーバに発行される予定である。

f) 新規S T PネットワークI D有効(S V)1310：このフィールドは、新規S T PネットワークI Dフィールドが有効であるかどうかを示す。それが有効ではない場合、C T N I DのS T PネットワークI D部分はそのコマンドによって変更されない。 40

g) 新規E T RネットワークI D有効(E V)1312：このフィールドは、新規E T RネットワークI Dフィールドが有効であるかどうかを示す。それが有効ではない場合、C T N I DのE T RネットワークI D部分はそのコマンドによって変更されない。

h) 新規S T PネットワークI D1314：S Vインジケータが有効である場合、このフィールドは新規S T PネットワークI Dを含む。

i) 新規E T RネットワークI D1316：E Vインジケータが有効である場合、このフィールドは新規E T RネットワークI Dを含む。

j) E T Rポート0マニュアル・ポート状態1318：E Vインジケータが有効であ

10

20

30

40

50

る場合、このフィールドは、ETRポート0の意図された状態を指定するマニュアル・ポート状態コードを含む。

k) ETRポート1マニュアル・ポート状態1320: EVインジケータが有効である場合、このフィールドは、ETRポート1の意図された状態を指定するマニュアル・ポート状態コードを含む。

#### 【0065】

図17に関連して、CTN ID変更コマンド用のコマンド応答ブロック1350の一実施形態について説明する。一例では、CTN ID変更コマンド応答ブロック13500は、たとえば、以下のものを含む。

a) 長さ1352: このフィールドは、コマンド応答ブロック長を指定する値を含む 10

。

b) 応答コード1354: このフィールドは、コマンド用の応答コードを含む。

c) コマンド・トランザクション番号1356: このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供される値を含む。

#### 【0066】

一例では、CTN ID変更がローカルCTN更新として発行される場合、STP機能は直ちに更新を実行し、構成変更通知コマンドはコンソールに発行されない。さらに、CTN ID変更がグローバルCTN更新として発行される場合、その更新が最新のものになったときに、アクティブ階層1サーバが構成変更通知コマンドをコンソールに発行する 20

#### 【0067】

CTN ID変更がグローバルCTN更新として発行される場合、STP機能は、CTN内のサーバでCTN変更操作をスケジュールするためにCTNパラメータ更新手順を実行する。スケジュールされた更新時間は、スケジュールされた更新時間の前にCTN内のサーバが更新情報を受信することを保証する値に設定される。指定された応答コードは、その操作がCTN内で正常にスケジュールされたことを示す。

#### 【0068】

更新がサーバでスケジュールされる場合、そのサーバは、スケジュールされたパラメータ更新情報をコンソール・ディスク記憶装置上に保管する。さらに、パラメータ更新がサーバで最新のものになると、サーバは更新されたCTNパラメータをコンソール・ディスク記憶装置上に保管する 30

#### 【0069】

一例として、その情報を保管するために、CTNパラメータ保存コマンドというコマンドがサーバによって発行される。このコマンドは、たとえば、長さフィールドおよびコマンド・コード・フィールドならびにコマンド・トランザクション番号と、CII Bと、STP機能が使用可能になっているかどうかを示すSTP機能使用可能インジケータと、サーバがローカル・クロック・サーバとして指定されているかどうかを指定するローカル・クロック・ソース・インジケータと、保存コマンドの時点でアクティブである発振器を識別するアクティブ発振器インジケータと、NCII Bと、SCII Bと、階層1サーバとして動作しているときのサーバ用のベース・ステアリング・レートを含む階層1ベースのステアリング・レートと、NSCI Bと、PCSI Bと、TCPBI Bと、LSOI Bと、保存コマンドが発行された時間を示す保存タイムスタンプと、発振器識別子0と、算出発振器スキー0と、発振器識別子1と、算出発振器スキー1の各フィールドを有する要求ブロックを含む。図19に関連して、これらのフィールドの多くについて以下に説明する。このコマンド用のコマンド応答ブロックは、たとえば、長さと、応答コードと、コマンド・トランザクション番号とを含む。 40

#### 【0070】

ローカルCTN ID変更は、ローカル・コマンドとして発行されたCTN ID変更コンソール・コマンドの結果としてサーバで行われる。CTN ID変更コンソール・コマンドを受信したサーバは、一例では、サーバでCTN IDに対して指定の変更を直ち 50

に（一例では）行い、サーバ上のすべての確立されたS T P経路を未初期化状態にし、次に、S T P長に関する初期化を実行するためのイニシアチブ（initiative）を確立する。C T N I D変更が行われると、構成変更マシン・チェック条件が生成される。

#### 【0071】

グローバルC T N I D更新は、新規C T N I Dブロックに指定されたC T N I D更新時間にサーバで行われる。サーバは、指定の更新時間にそのC T N I Dに対する変更を行う。変更後、サーバは、同期チェックしきい値（sync check threshold）に等しい期間の間、新規C T N I D値と前のC T N I D値との不一致の結果として発生するC T N I D不一致エラーを認識しない。新規C T N I D値と旧C T N I D値との不一致以外の不一致の結果として発生したC T N I D不一致エラーは、この期間の間、無視されない。C T N I D変更が行われると、構成変更マシン・チェック条件が生成される。

#### 【0072】

アクティブ階層1サーバは、グローバルC T N I D更新の結果として行われるC T N I D更新を行った後、構成変更通知コンソール・コマンドを発行する。

#### 【0073】

上述のコマンドに加えて、S T Pコンソール・インターフェースは、サーバからS T Pタイミングおよび構成情報を読み取るためのコマンドを含む。読み取りコマンドで提供される情報は、コンソール・ディスプレイを介して使用するために入手可能になる。図18に関連して、C T Nパラメータ読み取りコマンド用のコマンド要求ブロック1400の一例について説明する。一例では、C T Nパラメータ読み取りコマンド要求ブロック1400は、たとえば、以下のものを含む。

- a) 長さ1402：このフィールドは、コマンド・ブロック長を指定する値を含む。
- b) コマンド・コード1404：このフィールドは、C T Nパラメータ読み取りコマンドを指定する値を含む。
- c) コマンド・トランザクション番号1406：このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。

#### 【0074】

図19に関連して、C T Nパラメータ読み取りコマンド用のコマンド応答ブロック1450の一実施形態について説明する。一例では、応答ブロック1450は以下のものを含む。

- a) 長さ1452：このフィールドは、コマンド応答ブロック長を示すために使用される。
- b) 応答コード1454：このフィールドは、コマンド用の応答コードを含む。
- c) コマンド・トランザクション番号1456：このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供される値を含む。
- d) C T N I D情報ブロック1458：このフィールドは、サーバ用のC T N I D情報ブロック（C I I B）を含む。

一実施形態では、C T N I D情報ブロックは、サーバ用のC T N I Dと、非ヌルE T RネットワークにおけるE T Rポート0および1の状態を指定するコードとを含む。

e) 最大S T Pバージョン1460：このフィールドは、サーバによってサポートされる最大S T Pバージョン番号を示す値を含む。

f) アクティブS T Pバージョン1462：このフィールドは、サーバで現在アクティブであるS T Pバージョン番号を示す値を含む。

g) 最大タイミング階層レベル1464：このフィールドは、サーバを設定でき、サーバが同期状態になることができる最大階層レベルを指定する値を含む。最大タイミング階層レベルより大きい階層レベルを有するサーバは、非同期タイミング状態ならびに使用不能クロック・ソース状態にある。

h) 最大階層レベル1466：このフィールドは、C T N内の任意のサーバについて

10

20

30

40

50

設定できる最大階層レベルを指定する値を含む。

i ) ローカル・クロック・ソース ( L ) 1468 : このフィールドは、サーバがローカル・クロック・サーバとして指定されているかどうかを指定する。

j ) S T P クロック・ソース状態 ( C ) 1470 : このフィールドは、サーバに関するクロック・ソース状態を指定する。

k ) アクティブ発振器 ( A ) 1472 : このフィールドは、S T P 機能によってアクティブであると見なされる発振器を識別する。

l ) タイミング・モード ( T M D ) 1474 : このフィールドは、サーバのタイミング・モードを指定する。

m ) S T P タイミング状態 ( T S T ) 1476 : このフィールドは、サーバのタイミング状態を指定する。 10

n ) C T N タイプ ( C T N T ) 1478 : このフィールドは、サーバで構成されるタイミング・ネットワークのタイプを指定するコードを含む。タイプ例としては以下のものを含む。

定義済み C T N なし : サーバは、C T N への接続用に構成されていない。

S T P 専用タイミング・ネットワーク : サーバは、S T P ネットワークのみを含む C T N への接続用に構成されている。

混合タイミング・ネットワーク : サーバは、S T P ネットワークと E T R ネットワークの両方を含む C T N への接続用に構成されている。 20

o ) 階層 1480 : このフィールドは、サーバの階層レベルを示す値を含む。

p ) 新規 C T N I D 情報ブロック ( N C I I B ) 1482 : このフィールドは、サーバ用の新規 C I I B を含む。新規 C T N I D 更新時間は、U T C ( 協定世界時 ) フォーマットで保管される。

一例では、新規 C T N I D 情報ブロックは、新規 C T N I D と、E T R ポート 0 マニュアル・ポート状態と、E T R ポート 1 マニュアル・ポート状態と、たとえば、新規 C T N I D が最新のものになる予定である時間を示すタイムスタンプである新規 C T N I D 更新時間とを含む。

q ) 現行階層 1 構成情報ブロック ( S C I B ) 1484 : このフィールドは、サーバ用の現行階層 1 構成ブロックを含む。

r ) 新規階層 1 構成情報ブロック ( N S C I B ) 1486 : このフィールドは、サーバ用の新規階層 1 構成ブロックを含む。 30

s ) P R T ( 1 次基準時間 ) 訂正ステアリング情報ブロック ( P C S I B ) 1488 : このフィールドは、サーバ用の P C S I B を含む。

一例では、P R T 訂正ステアリング情報ブロックは以下のものを含む。

1 . P R T 訂正ステアリング・レート開始時刻 : このフィールドは、P R T 訂正ステアリングを開始する予定の時刻を示すタイムスタンプを含む。

2 . P R T ソース識別子 : このフィールドは 1 次基準時間ソース識別子を含む。

3 . コンソール分散 ( consoledispersion ) : このフィールドは、提供されたコンソール分散を含む。

4 . U T C 分散 : このフィールドは、提供された U T C 分散を含む。 40

5 . P R T オフセット : このフィールドは、提供された 1 次基準時間オフセットを含む。

6 . P R T タイムスタンプ : このフィールドは、提供された P R T タイムスタンプに対応するタイムスタンプを含む。

t ) 時間帯制御パラメータ情報ブロック ( T C P I B ) 1490 : このフィールドは、サーバ用の T C P I B を含む。新規 T Z I B 更新時間および新規 D S T オフセット更新時間は、U T C タイムスタンプ・フォーマットに変換される。

一例では、時間帯制御パラメータ情報ブロックは、たとえば、以下のものを含む。

1 . アクティブ現地時間コード : このフィールドは、現地時間設定が C T N について効力があるかどうかを識別し、効力がある場合にその値を入手する方法を識別する値 50

を含む。例としては以下のものを含む。

指定現地時間なし。時間帯オフセットとDSTオフセットはどちらも0である。

自動更新付きATZIB有効 - ATZIBは有効であり、CTN用の現地時間設定を指定する。TCPIB内のアクティブDSTオフセット・フィールド、新規DSTオフセット・フィールド、および新規DSTO更新時間フィールドは有効であり、ATZIB内のDSTオン・アルゴリズムおよびDSTオフ・アルゴリズムに基づいて設定される。TTOフィールドは有効ではない。

自動更新なしATZIB有効 - ATZIBは有効であり、CTN用の時間帯オフセットを指定する。ATZIBに指定されたDSTオン・アルゴリズムおよびDSTオフ・アルゴリズムはいずれも使用されない。アクティブDSTオフセット・フィールド、新規DSTオフセット・フィールド、および新規DSTO更新時間フィールドは有効である。DSTオフセット設定コンソール・コマンドは、新規DSTO更新時間を変更するため、ならびに更新が最新のものになったときに現地時間が標準時間またはDSTになる予定であるかどうかを指定するために使用される。TTOフィールドは有効ではない。

TTOフィールド有効 - TTOフィールドは有効であり、CTNの現地時間オフセットを指定し、時間帯オフセットと任意の夏時間オフセットの両方を含む。ATZIBフィールド、アクティブDSTオフセット・フィールド、新規DSTオフセット・フィールド、および新規DSTO更新時間フィールドは有効ではない。

2. 新規現地時間コード：このフィールドは、新規TZIBがCTNについて保留中であるかどうかを識別する値を含む。例としては以下のものを含む。

指定新規現地時間なし。

自動更新付きNTZIB有効 - NTZIBフィールドおよびNTZIB更新時間フィールドは有効であり、CTN用の新規時間設定を指定する。新規TZIBは、新規NTZIB更新時間に最新のものになる。アクティブDSTオフセット、新規DSTオフセット、および新規DSTO更新時間は、NTZIBが最新のものになったときに、NTZIBで提供されるDSTオン・アルゴリズムおよびDSTオフ・アルゴリズムに基づいて設定される。

自動更新なしNTZIB有効 - NTZIBフィールドおよびNTZIB更新時間フィールドは有効であり、新規TZIB更新時間に最新のものになる予定の時間帯オフセットおよびDSTオフセットを指定する。TZIBに指定されたDSTオン・アルゴリズムおよびDSTオフ・アルゴリズムは使用されない。

3. 新規TZIB DST(D)：このフィールドは、新規現地時間コードが、自動DST更新なしTZIB有効を表す値に等しいときに有効であり、NTZIBが最新のものになったときにCTN用の現地時間を夏時間に設定するべきかどうかまたはNTZIBが最新のものになったときにその現地時間を標準時間に設定するべきかどうかを示す。

4. 合計時間オフセット：このフィールドは、有効である場合、サーバで効力があるDSTオフセットおよび時間帯オフセットの結果としてSTP機能で効力がある合計時間オフセットを指定する。このフィールドは、アクティブ時間コードがTTOフィールドを指定したときに有効である。

5. アクティブ時間帯情報ブロック(ATZIB)：このフィールドは、サーバで現在効力がある時間帯情報ブロック(TZIB)を含む。

一例では、TZIBは以下のものを含む。

a a) 時間帯アルゴリズム番号：このフィールドは、このエントリ用の時間帯アルゴリズム番号を識別する。

b b) 時間帯オフセット：このフィールドは、UTCからの時間帯の差を示す値を含む。

c c) 夏時間(DST)オフセット：このフィールドは、夏時間が効力があるときに適用される予定のDSTオフセットを示す値を含む。DSTオフセット値がCT

10

20

30

40

50

Nについてアクティブになる予定の時刻および日付は、DSTオン・アルゴリズムによって指定される。サーバ用のDSTオフセットが0に設定される予定の時刻および日付は、DSTオフ・アルゴリズムによって指定される。

d d ) 標準時間名： このフィールドは、DSTオフセットが効力がないときの時間帯を識別する。

e e ) 夏時間名： このフィールドは、DSTオフセットが効力があるときの時間帯を識別する。

f f ) DSTオン時間アルゴリズム： このフィールドは、DSTオフセットが適用される予定の日付および時刻を計算するために使用すべきアルゴリズムを指定する。このオフセットは、自動DSTスケジューリングがそのサーバについて効力があるときに自動的に適用される。10

g g ) DSTオフ時間アルゴリズム： このフィールドは、DSTオフセットが除去される予定の日付および時刻を計算するために使用すべきアルゴリズムを指定する。このオフセットは、自動DSTスケジューリングがそのサーバについて効力があるときに自動的に除去される。

6 . 新規時間帯情報ブロック (NTZIB) : このフィールドは、有効である場合、NTZIB更新時間によって指定された時刻に始まる、ATZIBを置き換える予定の時間帯情報ブロックを含む。

7 . アクティブ夏時間オフセット (DSTO) : このフィールドは、サーバで現在効力があるDSTオフセットを示す値を含む。20

8 . 新規夏時間オフセット (NDSTO) : このフィールドは、新規DSTO更新タイムスタンプによって指定された時刻に始まる、アクティブDSTOを置き換える予定のオフセットを指定する値を含む。

9 . 新規TZIB更新時間 : このフィールドは、新規TZIBが効力を発する予定の時刻を指定するタイムスタンプを含む。

10 . 新規DSTO時間 : このフィールドは、新規DSTオフセットが効力を発する予定の時刻を指定するタイムスタンプを含む。

u ) うるう秒オフセット情報ブロック (LSOIB) 1492 : このフィールドは、サーバ用のLSOIBを含む。新規LSO更新時間は、UTCタイムスタンプ・フォーマットに変換される。30

一例では、うるう秒オフセット情報ブロックは、たとえば、以下のものを含む。

1 . アクティブうるう秒オフセット提供 (P) : このフィールドは、たとえば、1に設定されると、アクティブうるう秒オフセットがオペレータによって提供された値を含むことを示す。このフィールドが、たとえば、0に設定されると、アクティブうるう秒オフセットは、マシン初期設定値を含み、オペレータによって設定されていない。

2 . アクティブうるう秒オフセット (LSO) : このフィールドは、STP機能で現在効力があるうるう秒の数を示す値を含む。

3 . 新規うるう秒オフセット (LSO) : このフィールドは、LSO更新時間によって指定された時刻にアクティブになる予定のうるう秒の数を示す値を含む。

4 . 新規うるう秒オフセット (LSO) 更新時間 : このフィールドは、新規うるう秒オフセットが効力を発する予定の時刻を指定するタイムスタンプを含む。40

v ) 発振器識別子0 (1494) : このフィールドは、発振器0用の識別子を含む。

w ) 算出発振器スキューリング (COS0) 1496 : このフィールドは、サーバにインストールされた発振器0用の公称周波数に対する算出スキューリングを指定する値を含む。

x ) 発振器識別子1 (1498) : このフィールドは、発振器1用の識別子を含む。

y ) 算出発振器スキューリング (COS1) 1499 : このフィールドは、サーバにインストールされた発振器1用の公称周波数に対する算出スキューリングを指定する値を含む。

【0075】

上記の情報は、本発明の一態様では、協定タイミング・ネットワークのサーバの保全性を保証するために使用される。この情報は、ネットワークのサーバに伝搬される。ネット50

ワークのサーバは、構成を入手し、構成を理解し、同じ構成に従う。一実施形態では、サーバは、階層1構成を把握せずに事前指定量の時間（たとえば、フリーホール期間）より長く動作することはできない。サーバがネットワークに加入したいと希望する場合、そのサーバはこの情報を入手し使用する。この情報は、CTNの状態ならびに追加情報を提供する。

#### 【0076】

本発明の他の一態様により、アクティブ階層1サーバで障害が発生した場合にアクティブ階層1サーバとして引き継ぐことができるよう、協定タイミング・ネットワーク用の階層1構成の一部として代替サーバを構成するための機能が提供される。これにより、CTN内のサーバがアクティブ階層1サーバの障害にかかわらずタイミング・ネットワーク内で同期を維持または獲得できるように、CTNに関する単一障害点 (single point of failure) が防止される。

#### 【0077】

上述したように、1次および代替階層1サーバを指定する階層1構成をCTNについて定義することができる。1次サーバは、通常、CTN用のアクティブ階層1サーバとして動作し、バックアップとして動作する代替サーバは非アクティブ階層1サーバという。非アクティブ階層1サーバは、アクティブ階層1サーバの障害を検出すると、それがCTNのアクティブ階層1サーバとして引き継いだことを（たとえば、CTNパラメータ更新手順を介して）CTNに通知し、それ自体の時刻機構に基づいてCTNを駆動する。故障中の階層1サーバは、階層0に設定され、リカバリ後にネットワークに再加入することができるが、この実施形態では階層1サーバとして再加入することはできない。さらに、一実施形態では、1次サーバが階層2サーバとしてCTNに再加入する場合、そのサーバはバックアップ階層1サーバの役割を引き受ける。コンソールは、必要であれば、CTN用のアクティブ階層1サーバとしてその役割を再開するよう1次サーバに指示することができる。

#### 【0078】

図20に関連して、アクティブ階層1サーバにおける障害の検出およびリカバリの実行に関連するロジックの一実施形態について説明する。最初に、ステップ1500で、非アクティブ階層1サーバはアクティブ階層1サーバにおける障害を検出する。たとえば、階層1システム・チェック信号 (SCS) の受信、アクティブ階層1サーバがチェック停止状態またはパワーオフ状態に入ったと判断するためのコンソールとの通信を含むコンソール支援リカバリ、第3のサーバであるアービタ・サーバを使用して、アクティブ階層1サーバを投票で排除する (vote out) トライアッド・リカバリを含む、アクティブ階層1サーバにおける障害を検出するための様々なメカニズムが存在する。これらの障害検出メカニズムについては、さらに詳細に後述する。

#### 【0079】

ステップ1502で、アクティブ階層1障害を検出したことに応答して、非アクティブ階層1サーバは引き継ぎ手順を実行する。さらに、ステップ1504で、アクティブ階層1サーバがそれが障害状態に入ったことを検出した場合、ステップ1506で、引き渡し手順を実行し、CTN用のアクティブ階層1サーバとしてのその役割を断念する。

#### 【0080】

一例として、非アクティブ階層1サーバによって実行されるステップ1500および1502は、アクティブ階層1サーバによって実行されるステップ1504および1506と実質的に並行して実行することができる。

#### 【0081】

障害検出およびエラー・リカバリに関する詳細については、構成のタイプに基づく手順に関連して以下に説明する。たとえば、階層1サーバ・リカバリは、CTNのシステム1構成に代替階層1サーバが指定されている場合にのみ実行される。代替サーバは、デュアル・サーバ構成とトライアッド構成のいずれにも指定される。したがって、デュアル・サーバ階層1構成リカバリおよびトライアッド階層1構成リカバリの両方のためにリカバリ

10

20

30

40

50

手順が提供されるが、それぞれのリカバリについては以下に説明する。

【0082】

デュアル・サーバ階層1構成リカバリでは、デュアル・サーバ構成内の非アクティブ階層1サーバは、アクティブ階層1障害を検出したことに応答して、アクティブ階層1サーバの役割を引き継ぐ。図21に関連して、デュアル・サーバ階層1構成リカバリに関連するロジックの一実施形態について説明する。最初に、ステップ1600で、非アクティブ階層1サーバはアクティブ階層1サーバにおける障害を検出する。アクティブ階層1障害は、たとえば、コンソール支援リカバリ手順が実行され、アクティブ階層1障害が発生したことを示した場合、または階層1システム・チェック信号が認識された場合にデュアル・サーバ構成で検出されるが、それぞれの場合については以下に説明する。

10

【0083】

アクティブ階層1サーバ障害を検出したことに応答して、ステップ1602で、非アクティブ階層1サーバはアクティブ階層1引き継ぎ手順を実行し、CTN用のアクティブ階層1サーバの役割を引き継ぐ。これで、デュアル・サーバ階層1構成リカバリの処理を終了する。

【0084】

デュアル・サーバ階層1構成リカバリに加えて、トライアッド階層1構成リカバリが提供される。図22に関連して、トライアッド階層1構成リカバリに関連するロジックの一実施形態について説明する。最初に、ステップ1700で、トライアッド構成内の非アクティブ階層1サーバはアクティブ階層1障害を認識する。例として、非アクティブ階層1サーバは、トライアッド・リカバリ手順が実行され、アクティブ階層1障害が発生したことを示した場合、またはコンソール支援リカバリ手順が実行され、アクティブ階層1障害が発生したことを示した場合に、アクティブ階層1障害を認識する。

20

【0085】

アクティブ階層1障害を検出したことに応答して、ステップ1702で、非アクティブ階層1サーバはアクティブ階層1引き継ぎ手順を実行し、CTNのアクティブ階層1サーバの役割を引き継ぐ。さらに、ステップ1704で、非アクティブ階層1サーバおよびアービタ・サーバの両方への接続を失ったことを検出したことに応答して、アクティブ階層1サーバはアクティブ階層1サーバの役割を断念する。アクティブ階層1サーバは後述するアクティブ階層1引き渡し手順を実行し、CTNのアクティブ階層1サーバの役割を断念する。

30

【0086】

その上、ステップ1706で、たとえば、非アクティブ階層1サーバがアクティブ階層1サーバへの接続を失い、アービタへの接続が可能である場合、または非アクティブ階層1サーバがアクティブ階層1通信タイムアウトを認識し、アービタへの接続が可能である場合、あるいはその両方の場合に、非アクティブ階層1サーバはトライアッド・リカバリ手順を実行する。アクティブ階層1通信タイムアウトは、たとえば、CTN用のフリーホール間隔満了の2秒前に検出され、これにより、CTN内で同期チェック条件が検出される前に階層1引き継ぎを行うことができる。フリーホール間隔に関する詳細は、Carlson他により2006年8月30日に出願され、「Coordinated Timing Network Configuration Parameter Update Procedure」という名称の米国特許出願第11/468352号に記載されている。

40

【0087】

一実施形態では、照会1707で、トライアッド・リカバリが失敗した場合、ステップ1708で、トライアッド構成内の非アクティブ階層1サーバは、図21に関連して説明したデュアル・サーバ・リカバリを実行する。しかし、トライアッド・リカバリが成功した場合、デュアル・サーバ・リカバリは迂回される。これで、トライアッド階層1構成リカバリの処理を終了する。

【0088】

リカバリ中に実行される様々な手順に関する詳細については、さらに詳細に後述する。

50

## 【0089】

たとえば、アクティブ階層1サーバにおける障害を検出するためのメカニズムの1つは、非アクティブ階層1サーバで階層1システム・チェック信号(SCS)を受信することである。この階層1システム・チェック信号は、デュアル・サーバCTN構成内の非アクティブ階層1サーバに対し、CTN用のアクティブ階層1サーバとして動作し続けることを妨げるような状態にアクティブ階層1サーバが入ったことを示す。

## 【0090】

非アクティブ階層1サーバは、アクティブ階層1サーバが非アクティブ階層1サーバへのS TP接続性を終了したことを検出するときにSCSを認識する。非アクティブ階層1は、アクティブ階層1サーバに関連するS TP経路グループ内の最後の経路上でオフライン信号を受信するときに、アクティブ階層1サーバがS TP接続性を終了したことを検出する。

10

## 【0091】

アクティブ階層1サーバは、非アクティブ階層1サーバへのすべての接続性を終了するときに、それがSCSを発行したことを認識する。アクティブ階層1サーバは、非アクティブ階層1サーバに関連するS TP経路グループ内の最後の経路上でオフライン信号を送信するときに、非アクティブ階層1サーバへのすべてのS TP接続性が終了されたことを考慮する。

## 【0092】

障害を検出するためのもう1つのメカニズムはコンソール支援リカバリ手順であり、これは、アクティブ階層1サーバの状態を判断するためにデュアル・サーバ構成またはトライアッド構成のいずれかの非アクティブ階層1サーバによって開始される。この手順は、コンソールによって提供された情報を使用して、アクティブ階層1障害条件が存在するかどうかを判断する。

20

## 【0093】

図23に関連して、コンソール支援リカバリ手順に関連するロジックの一実施形態について説明する。最初に、ステップ1800で、非アクティブ階層1サーバがサーバ状態を要求する。特に、非アクティブ階層1サーバは、コマンド要求ブロック内のノード記述子フィールドがアクティブ階層1サーバのノード記述子に等しいS TP生成サーバ状態要求コンソール・コマンドを発行する。コンソールがこの要求を受信したことに応答して、ステップ1802で、コンソールは、アクティブ階層1サーバの状態を判断するためにアクティブ階層1サーバと通信しようと試みる。

30

## 【0094】

アクティブ階層1サーバの状態を判断しようと試みた後、ステップ1804で、コンソールは、要求ブロック内の状態コードがアクティブ階層1サーバの最も良く知られた状態に設定されたサーバ状態書き込みコンソール・コマンドを非アクティブ階層1サーバに発行する。ステップ1806で、非アクティブ階層1サーバは、その階層1サーバがチェック停止状態またはパワーオフ状態にあることを示すサーバ状態書き込みコンソール・コマンドを受信した場合、アクティブ階層1条件を認識する。

## 【0095】

40

図24に描写された構成に関連して、コンソール支援リカバリの一例に関連する詳細について説明する。一例として、デュアル・サーバ構成では、アクティブ階層1サーバに対する付加サーバ通信エラーを検出したときに、非アクティブ階層1サーバによってコンソール支援リカバリが実行される。以下の例では、アクティブ階層1サーバはチェック状態に入っている。図24の構成に関するコンソール支援リカバリ中に以下のアクションが実行される。

1. サーバD(1850)はサーバA(1852)への接続の喪失を検出する。
2. サーバDは、サーバ状態要求コマンドをコンソールに発行し、サーバAの状態を要求する。
3. コンソールは、その動作状態を判断するよう、サーバAに照会する。

50

4. コンソールは、サーバAが動作可能ではない（チェック停止またはパワーオフ）か、または動作可能状態が不明であることを示すサーバ状態書き込みをサーバDに発行する。

5. サーバAが動作可能であるかまたはその状態が不明である場合、サーバDはサーバ状態要求コマンドを繰り返す。

6. サーバAの状態が動作可能ではない場合、サーバDはCTNパラメータ更新手順を実行し、それ自体をアクティブ階層1サーバとして設定するよう新規階層1構成をスケジュールする。

一例では、協定タイミング・ネットワーク・パラメータ更新手順はタイミング・ネットワーク内で同時にCTNパラメータを更新する。一例として、この手順は、変更するための1つまたは複数のタイミング・パラメータを選択することと、変更すべきパラメータならびに変更が行われる予定の時期を含む情報応答パケットを構築することを含む。その後、そのパケットはネットワーク内の他のサーバにブロードキャストされる。受信側サーバでは、クロック・ソースとの通信が失われたかどうかが判断される。これは、典型的には、クロック・ソースとして使用されるネットワーク内のノードからタイミング・メッセージを定期的に受信することによって達成される。通信喪失の持続が長すぎたと判断された場合（これは、そのネットワークについて定刻より進んでいると定義され、一例として、サーバのクロック・ソースのクロックに対するサーバのクロックの最大ドリフト・レートの関数である）、受信側サーバはそのタイミング・パラメータを無効であると宣言する。続行するために、サーバは新規タイミング・パラメータを取得する。それが長すぎなかった場合、タイミング・パラメータは依然として有効であり、そのパラメータはネットワーク内のすべてのサーバにおいて推奨された将来の時点で更新される。

7. 階層1構成変更が効力を発すると、サーバDは構成変更通知をコンソールに発行する。

8. サーバC（1854）は階層3に変化し、サーバF（1856）は階層2に変化する。

9. サーバCおよびFに関する階層1構成変更および階層レベル変更を報告するために、ネットワークのすべてのサーバで構成変更マシン・チェックが生成される。

#### 【0096】

実行されるもう1つのリカバリ手順はトライアッド・リカバリ手順である。トライアッド・リカバリ手順は、階層1障害が発生したかどうかを判断するためにトライアッド構成内の非アクティブ階層1サーバによって実行される。

#### 【0097】

図25に関連して、トライアッド・リカバリ手順に関連するロジックの一実施形態について説明する。最初に、ステップ1900で、非アクティブ階層1サーバは引き継ぎモード状態に入る。さらに、ステップ1902で、アービタ引き継ぎモード設定メッセージ・コマンドがアービタ・サーバに発行される。アービタ引き継ぎモード設定操作は、たとえば、アービタを引き継ぎモードにするために、代替階層1サーバによりアービタ・サーバへのメッセージを介して発行される。受信側サーバは、アービタ引き継ぎ状態フラグおよび応答ブロックを返す。メッセージ・コマンド・データ・フィールドでは、アクティブ階層1通信タイムアウト・フラグは以下のように設定される。すなわち、サーバがアクティブ階層サーバへの接続が可能ではない場合、フラグはたとえば0に設定され、サーバがアクティブ階層サーバへの接続が可能であり、アクティブ階層1通信が認識された場合、フラグはたとえば1に設定される。

#### 【0098】

アービタ引き継ぎモード設定メッセージ・コマンドに対する応答が1に等しい引き継ぎ状態フラグを有する場合、非アクティブ階層1サーバはアクティブ階層1障害を認識し、アービタ引き継ぎモード設定メッセージ・コマンドに対する応答が0に等しい引き継ぎ状態フラグを有する場合、非アクティブ階層1サーバはアクティブ階層1障害を認識せず、以下のように実行する。すなわち、その応答が1に等しいアクティブ階層1通信タイムア

10

20

30

40

50

ウト・フラグを有する場合、サーバは引き継ぎモードを終了し、その応答が 0 に等しいアクティブ階層 1 通信タイムアウト・フラグを有する場合、非アクティブ階層 1 は引き継ぎモードのままになる。

【 0 0 9 9 】

非アクティブ階層 1 サーバは、以下のいずれかが発生したときに、引き継ぎモードを離れる。

0 に等しい引き継ぎ状態フラグと、1 に等しい通信タイムアウト・フラグとを含む、アービタ引き継ぎ設定メッセージ応答を受信する。

アービタ引き継ぎモード・リセット・コマンドをアービタに発行する。アービタ引き継ぎモード・リセット操作は、引き継ぎモードからアービタを取り出すために代替階層 1 サーバによってアービタ・サーバに発行される。この操作は、メッセージを介してアービタ・サーバに転送される。10

アクティブ階層 1 引き継ぎ手順を実行する。

【 0 1 0 0 】

非アクティブ階層 1 サーバは、引き継ぎモードにある間にアクティブ階層 1 サーバから確立された S T P 経路メッセージ・コマンドを受信した場合、アービタ引き継ぎモード・リセット・コマンドをアービタに発行し、アクティブ階層 1 サーバへの経路上で S T P 経路初期化を実行するためのイニシアチブを確立する。

【 0 1 0 1 】

アービタ・サーバは非アクティブ階層 1 サーバからアービタ引き継ぎモード設定メッセージ・コマンドを受信すると、アービタ引き継ぎモードに入る。アービタ・サーバは、以下のいずれかが発生したときに、アービタ引き継ぎモードを離れる。20

0 に等しい引き継ぎ状態フラグと、1 に等しい通信タイムアウト・フラグとを含む応答により、アービタがアービタ引き継ぎ設定メッセージ・コマンドに応答する。

アービタが非アクティブ階層 1 サーバからのアービタ引き継ぎモード・リセット・コマンドを受け入れる。

アービタが階層 1 構成更新を受信する。

【 0 1 0 2 】

アービタ引き継ぎモードにある間に、アービタは、引き継ぎ保留状態または引き継ぎアクティブ状態というアービタ引き継ぎ状態のうちの 1 つになる。このアービタ引き継ぎ状態は、アービタ引き継ぎモード設定コマンドに対する応答に示される。30

【 0 1 0 3 】

アービタは、アービタ引き継ぎモードに入り、以下の条件がいずれも存在するときに、引き継ぎ保留状態に入る。すなわち、アービタがアクティブ階層 1 サーバに接続されていることと、アクティブ階層 1 通信タイムアウトがアービタ内に存在しないことである。

【 0 1 0 4 】

引き継ぎ保留状態は、非アクティブ階層 1 サーバがアクティブ階層 1 引き継ぎを実行できないことを示す。アービタは、それが引き継ぎアクティブ状態に入ったときまたはアービタがアービタ引き継ぎモードを離れたときに、引き継ぎ保留状態を離れる。

【 0 1 0 5 】

アービタは、以下の状況の下で引き継ぎアクティブ状態に入る。

アービタがアービタ引き継ぎモードに入り、以下の条件のいずれかが存在する場合。すなわち、アクティブ階層 1 サーバがアービタに接続されていないこと、またはアクティブ階層 1 サーバがアービタに接続されており、アクティブ階層 1 通信タイムアウト条件がアービタで認識されたこと。

アービタが引き継ぎ保留状態にあり、アクティブ階層 1 サーバが接続状態を離れたか、またはアクティブ階層 1 通信タイムアウトが認識された場合。

【 0 1 0 6 】

引き継ぎアクティブ状態は、それがアクティブ階層 1 引き継ぎを実行できることを非アクティブ階層 1 サーバに示すものである。アービタがアービタ引き継ぎ設定コマンドに応50

答した後に引き継ぎアクティブ状態に入った場合、アービタは非アクティブ階層1サーバにアービタ引き継ぎアクティブ・コマンドを発行し（すなわち、メッセージを送信し）、それが引き継ぎ保留状態から引き継ぎアクティブ状態に遷移したことを示す。

#### 【0107】

アービタが引き継ぎアクティブ状態になると、アクティブ階層1サーバ用の経路グループ内の残りの経路はいずれも、通信エラーを示す未初期化状態に入る。引き継ぎアクティブ状態にある間に、アービタは、アクティブ階層1サーバからのS T P経路確立（E S P）メッセージ・コマンドに対し、引き継ぎアクティブ状態応答コードで応答する。

#### 【0108】

アービタは、アービタ引き継ぎモードを離れると、引き継ぎ保留状態を離れる。

10

#### 【0109】

図26に関連して、トライアッド構成内の階層1リカバリの一例について説明する。

1. サーバ1（1950）はチェック停止状態に入る。

2. サーバ2（1952）はサーバ1への接続の喪失を検出する。

3. サーバ2は引き継ぎモード設定制御コマンドをアービタ・サーバに発行する。

4. アービタは、以下のように引き継ぎモード設定コマンドに対する応答として引き継ぎモードに入る。

アービタは、サーバ1への接続が可能ではない場合、非アクティブ階層1サーバがアクティブ階層1サーバの役割を引き継ぐことができることを示す引き継ぎアクティブ状態に入ったと応答する。引き継ぎアクティブ状態にあるときに、アービタ・サーバはアクティブ階層1サーバにより経路を確立できないようにする。

20

アービタは、サーバ1への接続が可能ではない場合、非アクティブ階層1サーバがアクティブ階層1サーバの役割を引き継ぐことができないことを示す引き継ぎ保留状態に入ったと応答する。アービタは、引き継ぎ保留状態にある間にサーバ1への接続を失った場合、アービタ引き継ぎ状態アクティブ通知をサーバ2に発行する。

5. アービタは、引き継ぎモード設定コマンドまたはサーバ2へのアービタ引き継ぎ状態アクティブ通知コマンドのいずれかにより、それが引き継ぎアクティブ状態にあることを示す。

6. サーバ2は、C T Nパラメータ更新手順を実行し、それ自体をアクティブ階層1として設定する。

30

7. 階層1構成変更が最新ものになったときに、サーバ2は構成変更通知をコンソールに発行する。

8. 階層1構成変更を報告するために、ネットワークのすべてのサーバで構成変更マシン・チェックが生成される。

#### 【0110】

もう1つの例として、図27に関連して、アクティブS1へのリンクの喪失後の階層1サーバ引き継ぎのためのリカバリ手順について説明する。

1. 図27に関連して、サーバ2（1972）およびアービタ（1974）がサーバ1（1976）への接続の喪失を検出することが示されている。

2. サーバ2は、引き継ぎモード設定制御コマンドをアービタ・サーバに発行する。

40

3. アービタは、引き継ぎモード設定コマンドへの応答またはサーバ2へのアービタ引き継ぎ状態アクティブ通知コマンドのいずれかにより、それが引き継ぎアクティブ状態にあることを示し、それが引き継ぎを許可することを示す。

4. サーバ2は、C T Nパラメータ更新手順を実行し、それ自体をアクティブ階層1として設定する。

5. 階層1構成変更が最新ものになったときに、サーバ2は構成変更通知をコンソールに発行する。

6. サーバ1は、サーバ2および3への接続の喪失を検出し、引き渡し手順を実行する。そのサーバは、サーバBへの接続の結果として、階層3に降下する。

7. 階層1構成変更ならびにサーバ1が階層3にあることを報告するために、ネットワ

50

ークのすべてのサーバで構成変更マシン・チェック割り込み要求が生成される。

#### 【0111】

以下の例では、階層1引き継ぎは実行されない。図28に関連して、これについて説明する。この例では、以下のようになる。

1. サーバ2(1980)は、サーバ1(1982)への接続の喪失を検出する。
2. サーバ2は、引き継ぎモード設定制御コマンドをアービタ・サーバ(1984)に発行する。

3. アービタは、引き継ぎモード設定コマンドへの応答としてそれが引き継ぎ保留状態にあることを示し、非アクティブ階層1サーバがアクティブ階層1サーバの役割を引き継ぐことができないことを示す。

4. サーバ2は、リカバリを実行せず、階層3に降下する。
5. リンクが復元された場合、サーバ2は、引き継ぎモード・リセット・コマンドをアービタに発行する。
6. サーバ2は、階層レベル変更を報告するために構成変更マシン・チェック要求を生成する。

#### 【0112】

デュアル階層1リカバリ(図21)およびトライアッド階層1リカバリ(図22)のいずれでも、アクティブ階層1障害を認識したことに応答して、非アクティブ階層1サーバによってアクティブ階層1引き継ぎ手順が実行される。この手順の結果、非アクティブ階層1サーバはCTN用のアクティブ階層1サーバの役割を引き継ぐことになる。

#### 【0113】

図29に関連して、アクティブ階層1引き継ぎに関するロジックの一実施形態について説明する。最初に、ステップ2000で、アクティブ階層1サーバ用の経路グループ内のすべての残りの経路は未初期化状態になり、URCは通信エラーを示す。

#### 【0114】

さらに、ステップ2002で、サーバはその階層レベルを1に設定し、ステップ2004で、サーバおよびコンソールの階層1構成情報ブロックが変更される。一例では、そのサーバをアクティブ階層1サーバとして示すようにアクティブ・ビットが変更され、階層1最大短期スキューレート変更フィールドがそのサーバ用の値に設定され、階層1構成タイムスタンプが現行時間に設定され、他のフィールドが未変更になるように、SCIBが変更される。

#### 【0115】

上記に加えて、ステップ2006で、サーバはCTNパラメータ更新手順を実行し、CTN用の階層1構成情報ブロックを更新されたSCIB内の値に変更する。階層1構成更新時間および新規階層1構成情報ブロックはSCIB内の階層1構成タイムスタンプに設定される。これで、アクティブ階層1引き継ぎを終了する。

#### 【0116】

実行されるさらにもう1つの手順はアクティブ階層1引き渡し手順であり、これは、アクティブ階層1サーバ障害状態に入ったことを認識したときにアクティブ階層1サーバによりデュアル・サーバ構成またはトライアッド構成で実行される。この手順の結果、アクティブ階層1サーバはCTN用のアクティブ階層1サーバの役割を断念する。

#### 【0117】

図30に関連して、アクティブ階層1引き渡し手順に関するロジックの一実施形態について説明する。一例では、ステップ2100で、アクティブ階層1サーバにおける非アクティブ階層1およびアービタ・サーバ用の経路および経路グループが未初期化状態に設定され、URCは初期化が完了していないことを示す。さらに、ステップ2102で、アクティブ階層1サーバはその階層レベルを0に設定し、2次サーバになる。ステップ2104で、2次サーバとしての前のアクティブ階層1サーバがその階層1構成をヌル構成に設定する。さらに、ステップ2106で、このサーバはそのリンクを介して潜在的クロック・ソースを突き止めようと試みる。これで、引き渡しプロセスを完了する。

10

20

30

40

50

## 【0118】

上述したように、サーバ状態要求コマンドは、要求ブロック内のノード記述子によって指定されたサーバの動作状況を要求するために使用される。指定のサーバの状況は、サーバ状態書き込みコマンドを使用して、コンソールによって非同期的に提供される。図31に関連して、サーバ状態要求コマンド用のコマンド要求ブロックの一実施形態について説明する。

## 【0119】

サーバ状態要求要求ブロック2200は、たとえば、以下のものを含む。

- a) 長さ2202：このフィールドは、コマンド・ブロック長を指定する値を含む。
- b) コマンド・コード2204：このフィールドはサーバ状態読み取りコマンドを指定する。

c) コマンド・トランザクション番号2206：このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。

d) ノード記述子2208：このフィールドは、状況が要求されているサーバのノード記述子を含む。

## 【0120】

図32に関連して、サーバ状態要求コマンド用のコマンド応答ブロック2250の一実施形態について説明する。応答ブロック2250は、たとえば、以下のものを含む。

- a) 長さ2252：このフィールドは、コマンド・ブロック長を指定する値を含む。
- b) 応答コード2254：このフィールドはコマンド用の応答コードを含む。
- c) コマンド・トランザクション番号2256：このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供された値を含む。

## 【0121】

サーバ状態書き込みコマンドは、要求ブロックに指定されたサーバの動作状況を提供する。このコマンドは、サーバ状態要求コマンドを受け入れた後でコンソールによって発行される。図33に関連して、サーバ状態書き込みコマンド用のコマンド要求ブロックの一実施形態について説明する。一例では、サーバ状態書き込み要求ブロック2300は以下のものを含む。

- a) 長さ2302：このフィールドは、コマンド・ブロック長を指定する値を含む。
- b) コマンド・コード2304：このフィールドは、サーバ状態書き込みコマンドを指定する値を含む。
- c) コマンド・トランザクション番号2306：このフィールドは、発行されているコマンドに関連する値を含む。コマンド・トランザクション番号は、そのコマンド用の応答ブロックで返される。
- d) 状態コード2308：このフィールドは、要求ブロック内のノード記述子によって指定されたサーバの状態を示す値を含む。状態としては、たとえば、以下のものを含む。

サーバは動作可能である。

サーバ状態は不明である。

サーバはチェック停止状態にある。

サーバはパワーオフされている。

- e) ノード記述子2310：このフィールドは、状況が報告されているサーバのノード記述子を含む。

## 【0122】

図34に関連して、サーバ状態書き込みコマンド用のコマンド応答ブロックの一実施形態について説明する。一例では、サーバ状態書き込み応答ブロック2350は以下のものを含む。

a) 長さ2352：このフィールドは、コマンド・ブロック長を指定する値を含む。

b) 応答コード2354：このフィールドはコマンド用の応答コードを含む。

10

20

30

40

50

c) コマンド・トランザクション番号 2356：このフィールドは、コマンド要求ブロックのコマンド・トランザクション番号フィールドで提供された値を含む。

【0123】

一実施形態では、本発明の1つまたは複数の態様は、ネイティブ・アーキテクチャと呼ぶことのできるあるアーキテクチャに基づくが、ゲスト・アーキテクチャと呼ぶこともできる他のアーキテクチャをエミュレートする処理環境で実行することができる。例として、ネイティブ・アーキテクチャは、ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社によって提供されるPower 4またはPower PC（商標）アーキテクチャあるいはインテル社によって提供されるIntel（商標）アーキテクチャであり、ゲスト・アーキテクチャは、同じくニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社によって提供されるz/Architecture（商標）である。z/Architecture（商標）の諸態様は、2005年9月発行のIBM資料No. SA22-7832-04「z/Architecture Principles of Operation」に記載されている。このような環境では、z/Architecture（商標）で指定され、z/Architecture（商標）マシン上で実行されるように設計された命令またはロジックあるいはその両方は、z/Architecture（商標）以外のアーキテクチャ上で実行されるようにエミュレートされる。図35～図36に関連して、この処理環境の一例について説明する。

【0124】

図35を参照して、本発明の1つまたは複数の態様を取り入れて使用するための処理環境の一実施形態について説明する。処理環境2400は、たとえば、ネイティブ中央演算処理装置2402と、メモリ2404（たとえば、メイン・メモリ）と、たとえば、1つまたは複数のバス2408またはその他の接続部あるいはこれらの組み合わせを介して相互に結合された1つまたは複数の入出力（I/O）装置2406とを含む。例として、処理環境2400は、ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社によって提供されるPower PC（商標）プロセッサ、pSeries（商標）サーバ、またはxSeries（商標）サーバ、カリフォルニア州パロアルトのヒューレットパッカード社によって提供されるIntel（商標）Itanium（商標）2プロセッサ付きHP Superdome、IBM（商標）、ヒューレットパッカード社、Intel（商標）、サン・マイクロシステムズ社、その他によって提供されるアーキテクチャに基づくその他のマシン、あるいはこれらの組み合わせを含むことができる。Power PC（商標）、pSeries（商標）、およびxSeries（商標）は米国ニューヨーク州アーモンクのインターナショナル・ビジネス・マシーンズ社の登録商標である。Intel（商標）およびItanium（商標）2はカリフォルニア州サンタクララのインテル社の登録商標である。

【0125】

ネイティブ中央演算処理装置2402は、この環境内での処理中に使用される、1つまたは複数の汎用レジスタあるいは1つまたは複数の特殊目的レジスタもしくはこれらの組み合わせなどの1つまたは複数のネイティブ・レジスタ2410を含む。これらのレジスタは、任意の特定の時点におけるこの環境の状態を表す情報を含む。

【0126】

その上、ネイティブ中央演算処理装置2402は、メモリ2404に保管された命令およびコードを実行する。特定の一例では、中央演算処理装置は、メモリ2404に保管されたエミュレータ・コード2412を実行する。このコードは、あるアーキテクチャで構成された処理環境で他のアーキテクチャをエミュレートできるようにするものである。たとえば、エミュレータ・コード2412は、Power PC（商標）プロセッサ、pSeries（商標）サーバ、xSeries（商標）サーバ、HP Superdome（商標）サーバ、その他などのz/Architecture以外のアーキテクチャに基づくマシンで、z/Architecture（商標）をエミュレートし、z/Architecture（商標）に基づいて開発されたソフトウェアおよび命令を実行できるよう

10

20

30

40

50

にする。

【0127】

図36に関連して、エミュレータ・コード2412に関する詳細について説明する。ゲスト命令2502は、ネイティブCPU2402のアーキテクチャ以外のアーキテクチャで実行されるように開発されたソフトウェア命令（たとえば、マシン命令）を含む。たとえば、ゲスト命令2502は、z/Architecture（商標）902上で実行されるように設計された可能性があるが、その代わりにネイティブCPU2402（たとえば、Intel（商標）Itanium（商標）2プロセッサである可能性がある）上でエミュレートされる。一例では、エミュレータ・コード2412は、メモリ2404から1つまたは複数のゲスト命令2502を入手し、任意選択で、入手した命令のローカル・バッファを行うための命令取り出しルーチン2500を含む。10

【0128】

エミュレータ・コード2412は、入手されたゲスト命令のタイプを決定し、ゲスト命令に対応する1つまたは複数のネイティブ命令2509を提供するための命令変換ルーチン2504をさらに含む。一例では、提供することは、たとえば、変換プロセス中に所与のゲスト命令に関する命令のネイティブ・ストリームを作成することを含む。これは、関数を識別することと、同等のネイティブ命令を作成することを含む。他の一例では、ネイティブ命令を提供することは、ゲスト命令に関連するコード・セグメントをエミュレータ内で選択することを含む。たとえば、各ゲスト命令は、エミュレータ内に関連のコード・セグメントを有し、これは1つまたは複数のネイティブ命令のシーケンスを含み、そのコード・セグメントは実行するために選択される。20

【0129】

エミュレータ・コード2412は、ネイティブ命令を実行させるためのエミュレーション制御ルーチン2506をさらに含む。エミュレーション制御ルーチン2506により、ネイティブCPU2402は、1つまたは複数の前に入手したゲスト命令をエミュレートするネイティブ命令のルーチンを実行し、このような実行の終わりに、命令取り出しルーチンに制御を返して次のゲスト命令またはゲスト命令のグループの入手をエミュレートすることができる。ネイティブ命令2509の実行は、メモリ2404からレジスタにデータをロードすること、レジスタからメモリにデータを戻して保管すること、または変換ルーチンによって決定されたように何らかのタイプの算術演算または論理演算を実行することを含むことができる。各ルーチンは、たとえば、メモリに保管され、ネイティブ中央演算処理装置2402によって実行されるソフトウェアで実現される。その他の例では、これらのルーチンまたは演算の1つまたは複数は、ファームウェア、ハードウェア、ソフトウェア、またはこれらの何らかの組み合わせで実現される。エミュレートされたゲスト・プロセッサのレジスタは、ネイティブCPUのレジスタ2410を使用するか、またはメモリ2404内の位置を使用することによってエミュレートすることができる。諸実施形態では、ゲスト命令2502、ネイティブ命令2509、およびエミュレーション・コード2412は、同じメモリに常駐するか、または異なるメモリ・デバイス間に分散させることができる。30

【0130】

さらに他の実施形態では、プログラム・コードの保管または実行あるいはその両方に適したデータ処理システムであって、直接またはシステム・バスを介して間接的にメモリ・エレメントに結合された少なくとも1つのプロセッサを含むものが使用可能である。メモリ・エレメントは、たとえば、プログラム・コードの実際の実行中に使用されるローカル・メモリ、大容量記憶装置、ならびに、実行中にコードを大容量記憶装置から取り出さなければならない回数を削減するために少なくとも何らかのプログラム・コードの一次記憶を行うキャッシュ・メモリを含む。40

【0131】

入出力またはI/O装置（キーボード、ディスプレイ、ポインティング・デバイス、DASD、テープ、CD、DVD、サム・ドライブ、およびその他の記憶媒体などを含むが50

、これらに限定されない)は、直接または介在する I / O コントローラを介してシステムに結合することができる。データ処理システムが介在する私設網または公衆網を介して他のデータ処理システムまたはリモート・プリンタあるいは記憶装置に結合された状態になれるようするために、ネットワーク・アダプタもシステムに結合することができる。モデム、ケーブル・モデム、およびイーサネット・カードは、使用可能なタイプのネットワーク・アダプタのうちのいくつかに過ぎない。

【 0 1 3 2 】

本発明の 1 つまたは複数の態様は、たとえば、コンピュータ使用可能媒体を有する装置 (article of manufacture) (たとえば、1 つまたは複数のコンピュータ・プログラム) に含めることができる。この媒体は、たとえば、本発明の諸機能を提供し促進するためのコンピュータ可読プログラム・コード手段またはロジック (たとえば、命令、コード、コマンドなど) をそこに有する。この装置は、システム (たとえば、コンピュータ・システム) の一部として含めるか、または別個に販売することができる。

【 0 1 3 3 】

図 3 7 に関連して、本発明の 1 つまたは複数の態様を取り入れた装置またはコンピュータ・プログラムの一例について説明する。コンピュータ・プログラム 2 6 0 0 は、たとえば、本発明の 1 つまたは複数の態様を提供し促進するためのコンピュータ可読プログラム・コード手段またはロジック 2 6 0 4 をそこに保管するための 1 つまたは複数のコンピュータ使用可能媒体 2 6 0 2 を含む。この媒体は、電子、磁気、光学、電磁、赤外線、または半導体のシステム (あるいは装置またはデバイス) もしくは伝搬媒体にすることができる。コンピュータ可読媒体の例としては、半導体またはソリッド・ステート・メモリ、磁気テープ、取り外し可能コンピュータ・ディスクケット、ランダム・アクセス・メモリ (R A M)、読み取り専用メモリ (R O M)、剛性磁気ディスク、および光ディスクを含む。光ディスクの例としては、コンパクト・ディスク読み取り専用メモリ (C D - R O M)、書き換え可能コンパクト・ディスク (C D - R / W)、および D V D を含む。

【 0 1 3 4 】

1 つまたは複数のコンピュータ可読プログラム・コード手段またはロジックによって定義された 1 つまたは複数の相關モジュールの論理アセンブリあるいは一連のプログラム命令は、本発明の 1 つまたは複数の態様の実行を指示するものである。

【 0 1 3 5 】

本明細書には、協定タイミング・ネットワークを形成するために複数の別個のコンピューティング・システムによる時刻同期の維持を促進する諸機能が記載されている。タイミング・ネットワーク内のサーバは、サーバ・タイム・プロトコルを使用して、既存の高速データ・リンクによりシステム間で計時情報を受け渡し、これにより、各システムの時刻機構を現代のハイエンド・コンピューティング・システムに必要な正確度で同期させるための機能を提供する。高速低待ち時間リンクによる S T P の使用は、單一サーバによって提供された基準時間に基づくときに、C T N 内のすべてのシステムを、たとえば、数マイクロ秒の正確度で同期させるための機能を提供する。

【 0 1 3 6 】

S T P は、時間帯、夏時間オフセット、うるう秒オフセットなど、C T N 内で計時情報を設定し維持するための機能を提供する。この情報は、すべての変更が C T N 内のすべてのサーバで同時に行われるように、計画的かつ一貫して C T N 内で更新することができる。これにより、これらのパラメータが無計画に更新されたときに発生し、コンピュータ間の時間設定の矛盾を生み出す潜在的なシステム露出や混乱が防止される。

【 0 1 3 7 】

C T N パラメータは、S T P コンソール・インターフェースを介してオペレータによって設定し読み取ることができる。C T N パラメータとしては、サーバ接続性、時間帯および夏時間などの現地時間情報、ならびに U T C を計算するために必要なうるう秒を含む。コンソール自体は、C T N パラメータを表示し設定するためのオペレータ・インターフェースを提供し、S T P 機能と通信するための機能を有する任意の要素である。

10

20

30

40

50

## 【0138】

本発明の一態様では、單一アクティブ階層1サーバは、ネットワーク用のクロック・ソースとして提供される。アクティブ階層1サーバの時刻機構は、任意の時刻に設定することができるが、一般に、UTCへのダイヤルアップなどの外部時間基準に設定される。その他のサーバは、ヌル構成として知られる未定義の階層1構成を有する場合、または1次タイム・サーバの階層1構成に一致する階層1構成を有する場合に、CTNに加入することができる。したがって、CTN内の同期正確度は、外部タイム・ソースの品質または階層1サーバの外部タイム・ソースの存在にさえ依存しない。これにより、CTN内のサーバが同じルート1次基準時間に同期することを保証する。

## 【0139】

本発明の他の一態様では、アクティブ階層1サーバで障害が発生した場合にアクティブ階層サーバとして引き継ぐことができるよう、代替サーバが定義される。これにより、CTN内のサーバがアクティブ階層1サーバの障害にかかわらずタイミング・ネットワーク内で同期を維持または回復できるよう、CTNに関する单一障害点が防止される。

## 【0140】

タイミング・ネットワークに関する追加情報は、「Server Time Protocol Messages And Methods」という名称の米国特許出願第60/887512号、「Channel Subsystem Server Time Protocol Commands」という名称の米国特許出願第60/887544号、「Method And System For Establishing A Logical Path Between Servers In A Coordinated Timing Network」という名称の米国特許出願第60/887576号、「Facilitating Synchronization Of Servers In A Coordinated Timing Network」という名称の米国特許出願第60/887584号、「Facilitating Recovery In A Coordinated Timing Network」という名称の米国特許出願第60/887586号、2006年8月30日に出願され、「Coordinated Timing Network Configuration Parameter Update Procedure」という名称の米国特許出願第11/468352号、2006年7月26日に出願され、「Directly Obtaining By Application Programs Information Usable In Determining Clock Accuracy」という名称の米国特許出願第11/460025号、「System And Method For TOD-Clock Steering」という名称の米国特許出願第11/223886号、「Synchronization Signal For TOD-Clock Steering Adjustment」という名称の米国特許出願第11/532168号、「Managing Data Access Via A LoopOnly If Changed Locking Facility」という名称の米国特許出願第11/468501号、「Clock Filter Dispersion」という名称の米国特許出願第11/223878号、「Method And System For Clock Skew And Offset Estimation」という名称の米国特許出願第11/223876号、「Use Of T4 Timestamps To Calculate Clock Offset And Skew」という名称の米国特許出願第11/223577号、および「System And Method For Calibrating A TOD Clock」という名称の米国特許出願第11/223642号という特許出願に示されている。

## 【0141】

有利なことに、本発明の1つまたは複数の態様は、同期および正確度要件を満たすことができる高品質の計時情報の提供を可能にする。CTN内のシステムは、UTCなど、外部タイム・ソースに対し、あまり厳しくない同期正確度の要件により、厳重に同期したままになる。高価な専用タイミング・リンクまたは個別外部ボックスあるいはその両方の使用は不要である。さらに、各サーバが外部タイム・サーバに接続するためまたはGPSを備えるための要件は不要である。

## 【0142】

本明細書には1つまたは複数の例が提供されているが、これらは例に過ぎない。本発明の精神を逸脱せずに、多くの変形例が可能である。たとえば、本明細書に提供されている例以外の処理環境は、本発明の1つまたは複数の態様を含むか、その態様から恩恵を受けるか、あるいはその両方である可能性がある。さらに、この環境は、z/Architecture(商標)に基づく必要はないが、その代わりに、たとえば、IBM(商標)、Intel(商標)、サン・マイクロシステムズ社、その他によって提供される他のア-

10

20

30

40

50

キテクチャに基づくものにすることができる。さらに、この環境は、例として、複数のプロセッサを含むか、区画化するか、または他のシステムに結合するか、あるいはこれらの組み合わせにすることができる。

【 0 1 4 3 】

その上、様々な制御ブロックが記載されているが、これらの制御ブロックのそれぞれは、追加情報を含むか、より少ない情報を含むか、または異なる情報を含むか、あるいはこれらの組み合わせになる可能性がある。制御ブロック内の位置および制御ブロック内の各フィールドのサイズは、種々の環境について様々になる可能性がある。

【 0 1 4 4 】

本明細書で使用する「入手する (obtaining)」という用語は、取り出すこと、受信すること、有すること、提供すること、提供されること、作成すること、開発することなどを含むが、これらに限定されない。 10

【 0 1 4 5 】

本明細書の1つまたは複数の態様の諸機能は、ソフトウェア、ファームウェア、ハードウェア、またはこれらの何らかの組み合わせで実現することができる。本発明の諸機能を実行するためにマシンによって実行可能な複数命令からなる少なくとも1つのプログラムを実施するマシンによって読み取り可能な少なくとも1つのプログラム記憶装置を提供することができる。

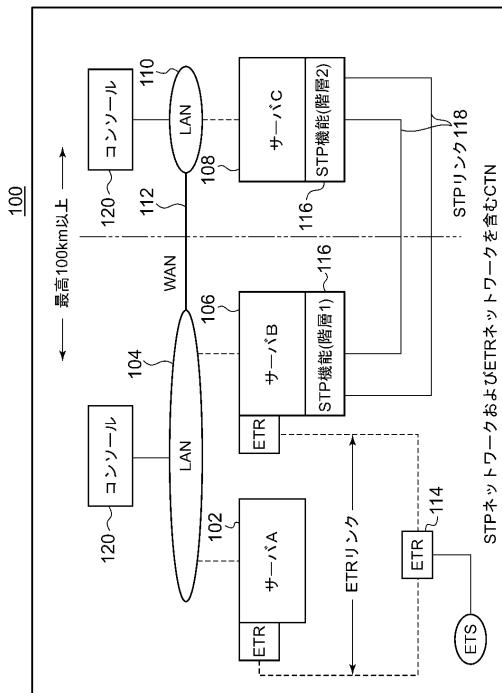
【 0 1 4 6 】

本明細書に描写されている流れ図は例に過ぎない。本発明の精神を逸脱せずに、これらの流れ図またはそこに記載されているステップ（または操作）に対する多くの変形例が可能である。たとえば、これらのステップを異なる順序で実行するか、ステップを追加、削除、または変更することができる。これらの変形例はいずれも、請求された本発明の一部と見なされる。 20

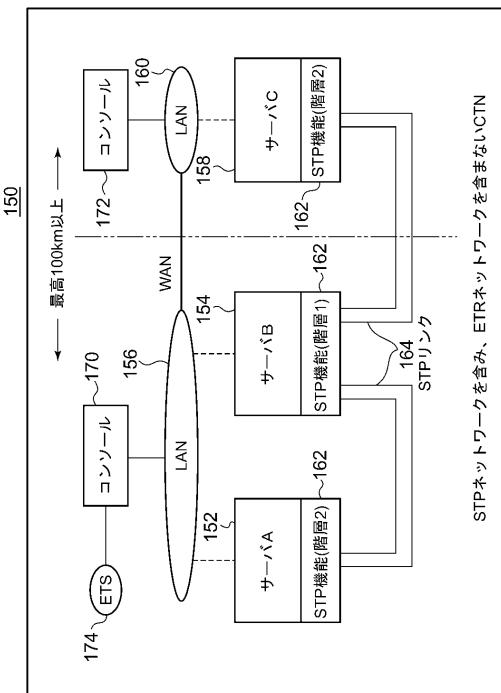
【 0 1 4 7 】

諸実施形態が本明細書に詳細に描写され記載されているが、本発明の精神を逸脱せずに、様々な変更、追加、置き換えなどが可能であることは当業者には明らかになることであり、したがって、これらは特許請求の範囲に定義された本発明の範囲内ものであると見なされる。

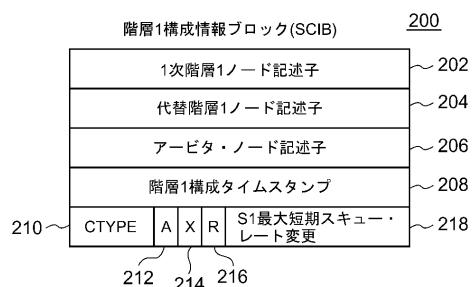
【図1】



【図2】



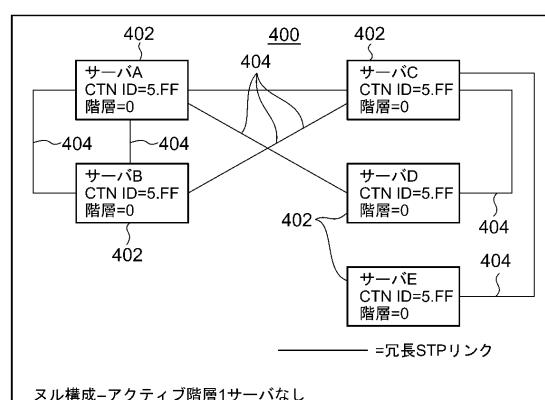
【図3】



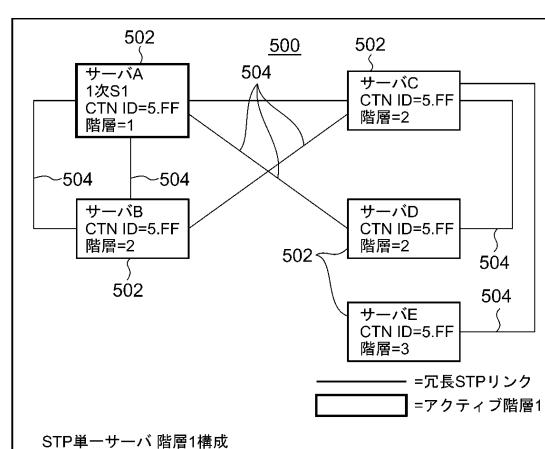
【図4】



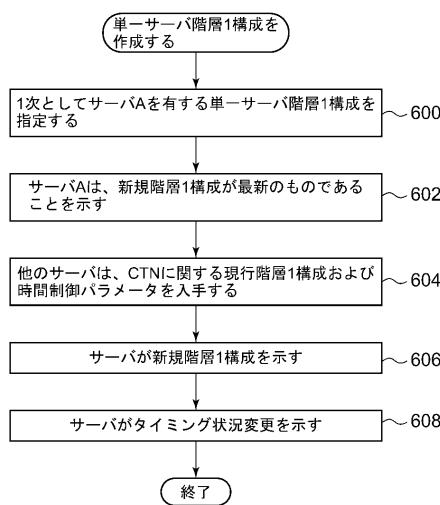
【図5】



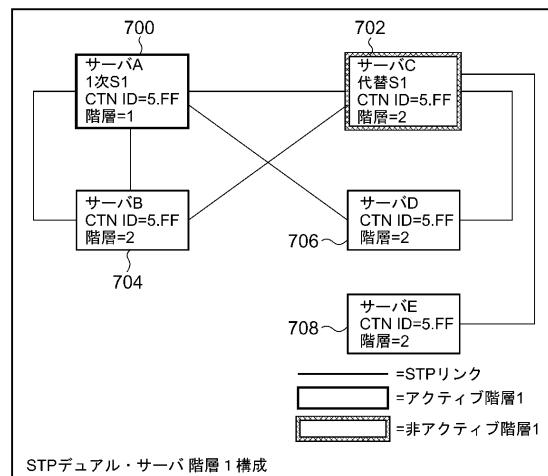
【図6】



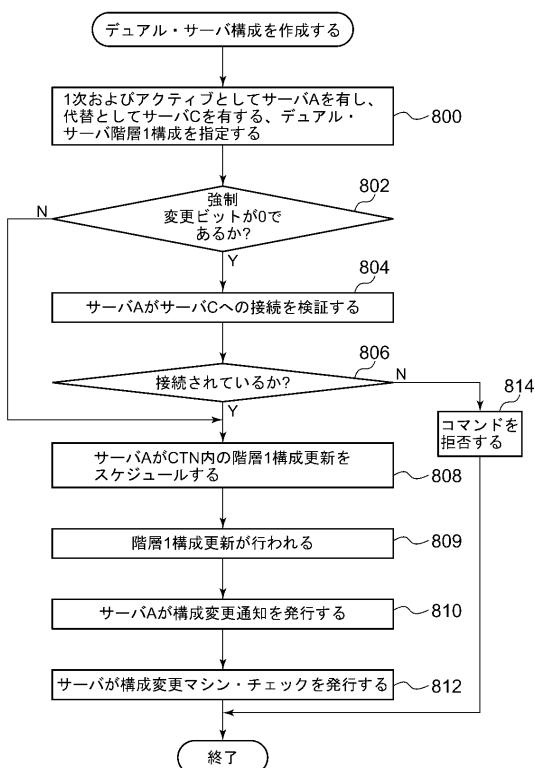
【図7】



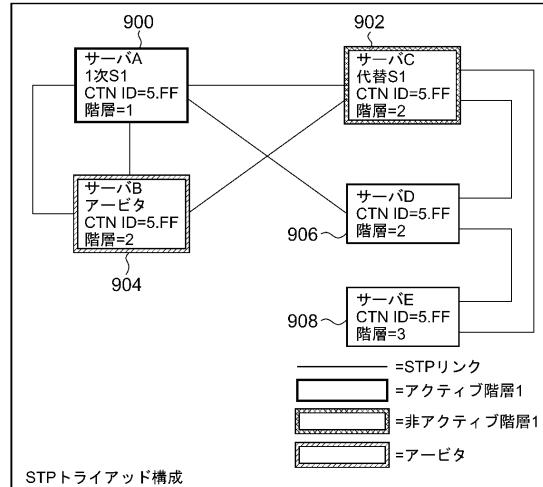
【図8】



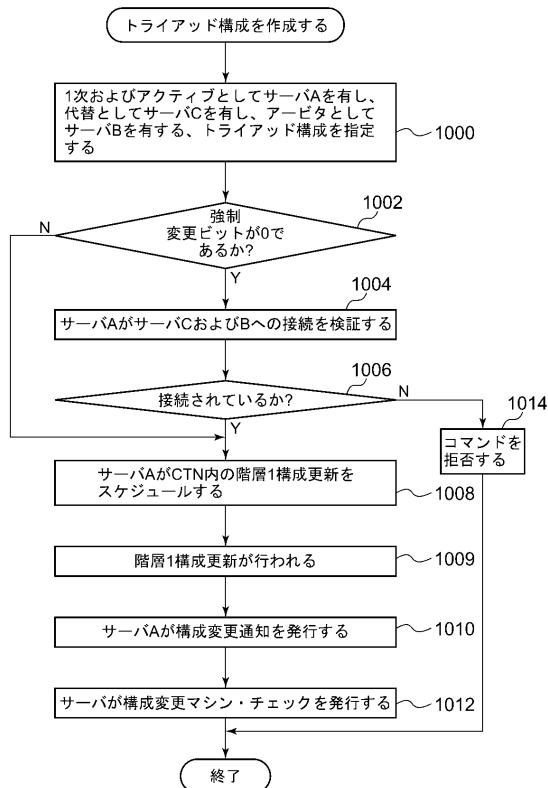
【図9】



【図10】



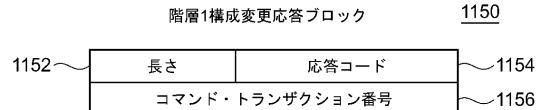
【図11】



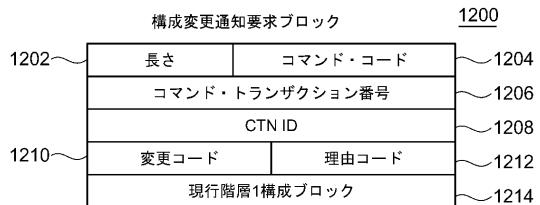
【図12】



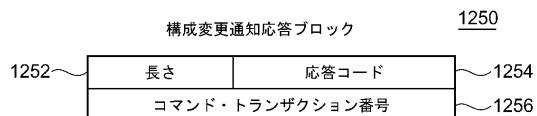
【図13】



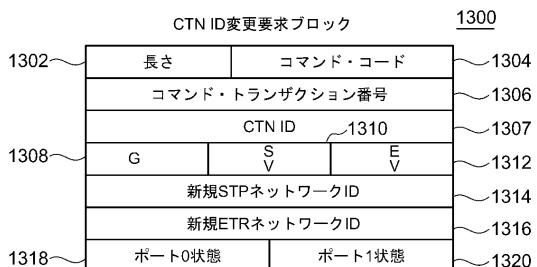
【図14】



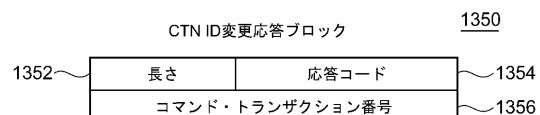
【図15】



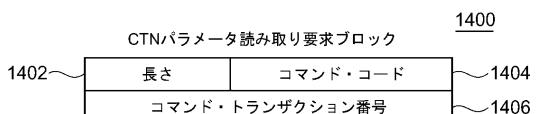
【図16】



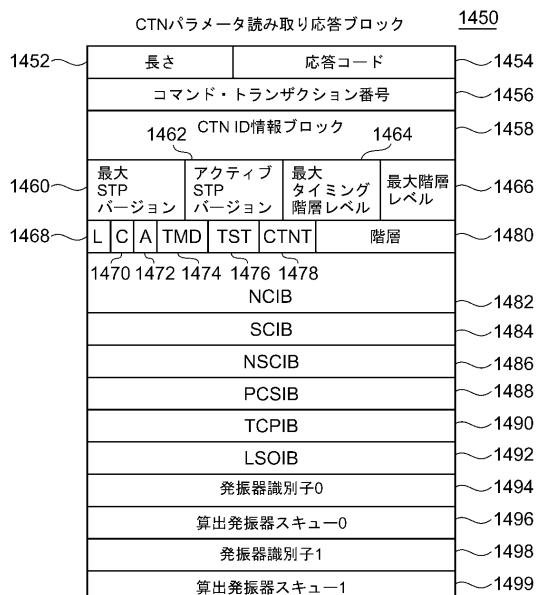
【図17】



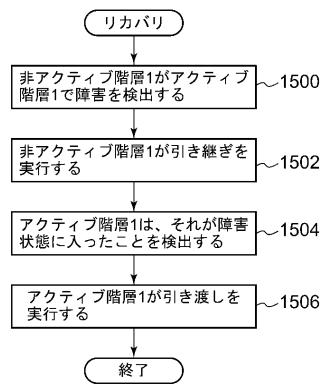
【図18】



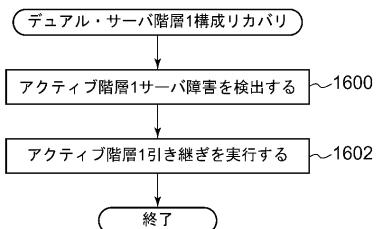
【図19】



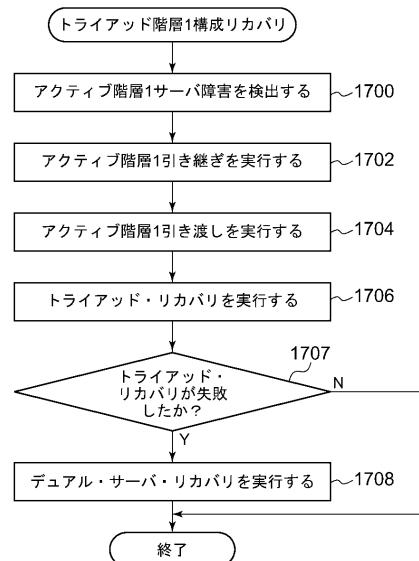
【図20】



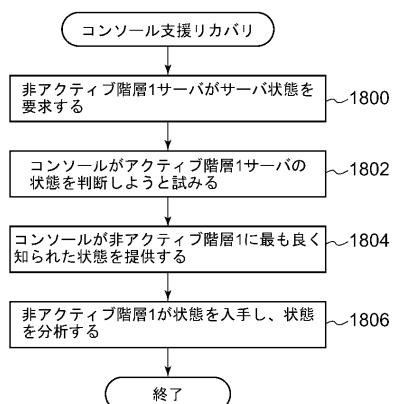
【図21】



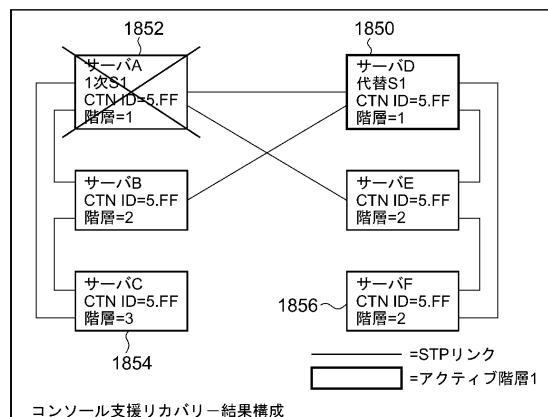
【図22】



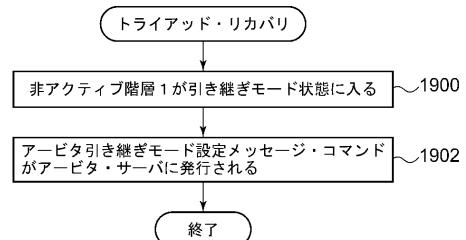
【図23】



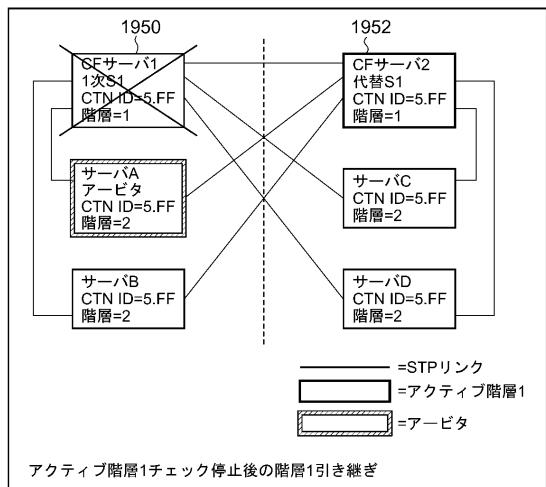
【図24】



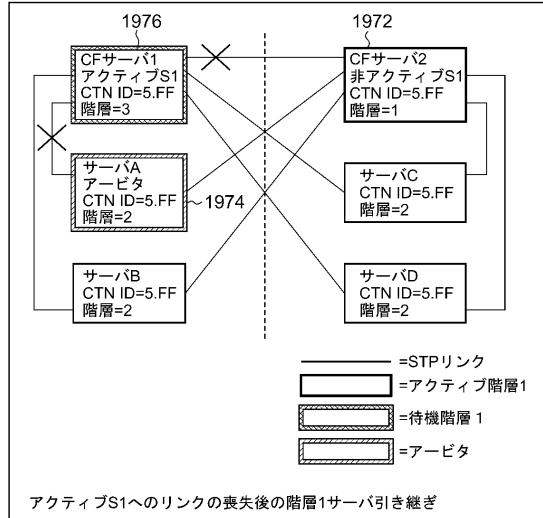
【図25】



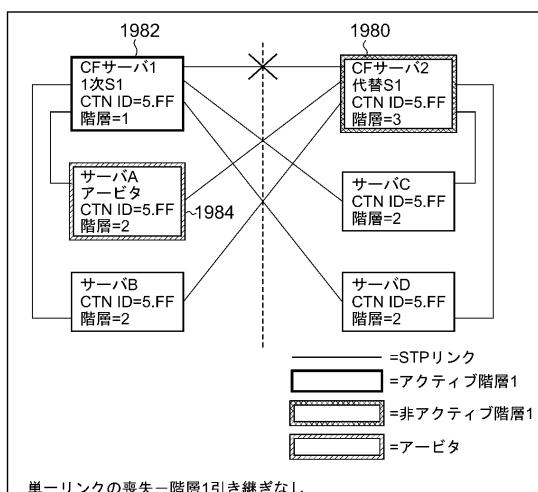
【図26】



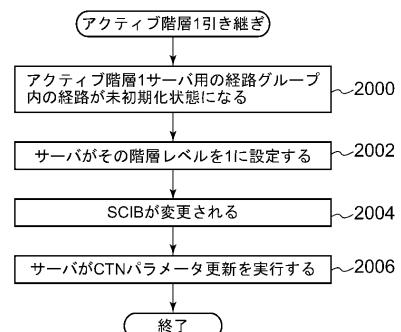
【図27】



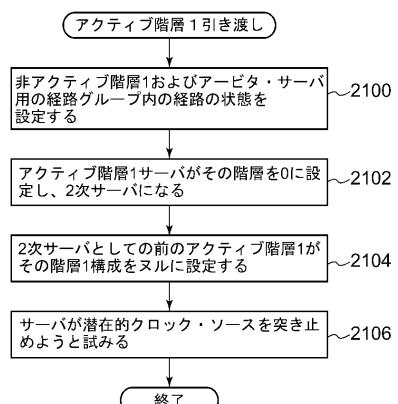
【図28】



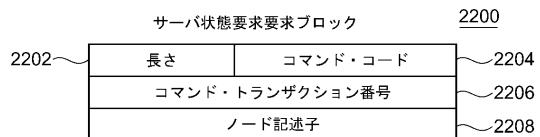
【図29】



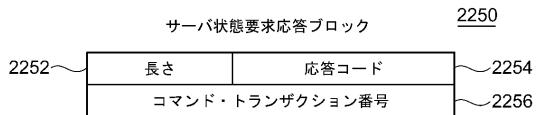
【図30】



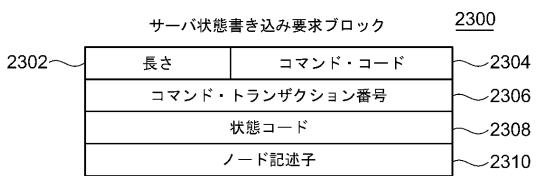
【図31】



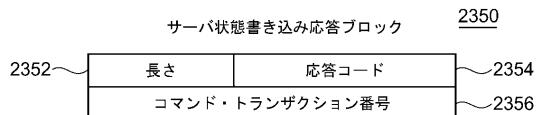
【図32】



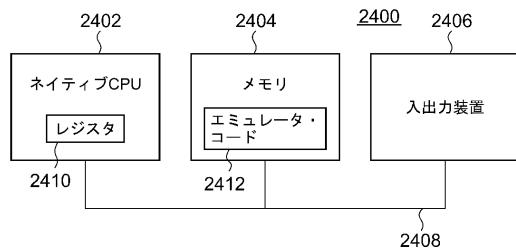
【図33】



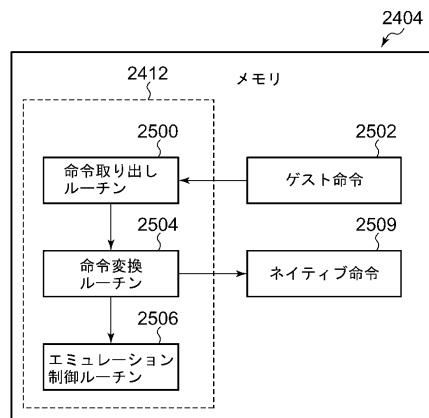
【図34】



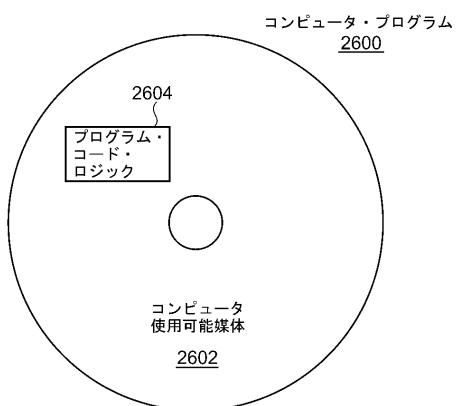
【図35】



【図36】



【図37】



---

フロントページの続き

- (72)発明者 カールソン、スコット  
アメリカ合衆国 85750 アリゾナ州ツーソン イースト・コルテ・デル・ジョーベン 8081
- (72)発明者 クラブツリー、ドナルド  
アメリカ合衆国 12466 ニューヨーク州ポート・エウェン セント・ジョセフ・ブルーヴァード 7
- (72)発明者 ダーレン、デニス  
アメリカ合衆国 12572 ニューヨーク州ラインベック ポンド・ドライヴ・ウエスト 23
- (72)発明者 セヴィニー、デニス  
アメリカ合衆国 12590 ニューヨーク州ワッピンガース・フォールズ レーン・ゲート・ロード 37
- (72)発明者 ウィアボウスキー、ジュディス  
アメリカ合衆国 13827 ニューヨーク州オウェゴ イースト・ビーチャー・ヒル・ロード 24  
89

審査官 鈴村 理絵子

- (56)参考文献 特開 2005-227192 (JP, A)  
特開 2000-155729 (JP, A)  
特開 2004-199366 (JP, A)  
特開平 11-184825 (JP, A)  
特開 2005-100030 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 13/00  
H04L 7/00