(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7429086号
(P7429086)

(45)発行日　令和6年2月7日(2024.2.7)

(24)登録日　令和6年1月30日(2024.1.30)

(51)国際特許分類

		FI		
G09C 1/00 (2006.01)		G09C	1/00	650Z
G06F 21/62 (2013.01)		G06F	21/62	318

請求項の数　8　（全32頁）

(21)出願番号　特願2021-563596(P2021-563596)
(86)(22)出願日　令和2年3月23日(2020.3.23)
(65)公表番号　特表2022-530947(P2022-530947A)
(43)公表日　令和4年7月5日(2022.7.5)
(86)国際出願番号　PCT/IB2020/052695
(87)国際公開番号　WO2020/222049
(87)国際公開日　令和2年11月5日(2020.11.5)
審査請求日　令和4年8月24日(2022.8.24)
(31)優先権主張番号　16/397,392
(32)優先日　平成31年4月29日(2019.4.29)
(33)優先権主張国・地域又は機関
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
New Orchard Road, Armonk, New York 10504, United States of America
(74)代理人　100112690
弁理士　太佐 種一

最終頁に続く

(54)【発明の名称】　分散による難読化に基づくセキュアなデータ・ストレージ

(57)【特許請求の範囲】

【請求項１】

データを安全に格納するための方法であって、

コンピュータ・プロセッサが、データを格納するための要求を受信することであって、前記データが複数の要素を含んでいる、前記受信することと、

前記コンピュータ・プロセッサが、前記データの前記複数の要素を各数値に変換することと、

前記コンピュータ・プロセッサが、１つまたは複数のエンコーディング・アルゴリズムを、変換された前記データに適用することによって、エンコードされたデータの複数の要素を生成することであって、前記エンコードされたデータの前記複数の要素の数が、変換された前記データの前記複数の要素の数に等しい、前記生成することと、

前記コンピュータ・プロセッサが、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散することと、

前記コンピュータ・プロセッサが、通信ユニットにより、前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信することと

を実行する、方法。

【請求項２】

データを安全に格納するための方法であって、

コンピュータ・プロセッサが、データを格納するための要求を受信することであって、

前記データが複数の要素を含んでいる、前記受信することと、
前記コンピュータ・プロセッサが、１つまたは複数のエンコーディング・アルゴリズムを前記データに適用することによって、エンコードされたデータの複数の要素を生成することであって、前記エンコードされたデータの前記複数の要素の数が、前記データの前記複数の要素の数に等しい、前記生成することと、
前記コンピュータ・プロセッサが、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散することと、
前記コンピュータ・プロセッサが、通信ユニットにより、前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信することと
を実行し、前記データをエンコードすることが、
前記コンピュータ・プロセッサが、前記データの前記複数の要素をエンコードするための１つまたは複数のエンコーディング・アルゴリズムを識別することをさらに含み、少なくとも１つのエンコーディング・アルゴリズムが、前記データの前記複数の要素のその後の要素をエンコードするために、前記データの第１の前にエンコードされた要素の利用を指示する、方法。
【請求項３】
データを安全に格納するための方法であって、
コンピュータ・プロセッサが、データを格納するための要求を受信することであって、前記データが複数の要素を含んでいる、前記受信することと、
前記コンピュータ・プロセッサが、１つまたは複数のエンコーディング・アルゴリズムを前記データに適用することによって、エンコードされたデータの複数の要素を生成することであって、前記エンコードされたデータの前記複数の要素の数が、前記データの前記複数の要素の数に等しい、前記生成することと、
前記コンピュータ・プロセッサが、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散することと、
前記コンピュータ・プロセッサが、通信ユニットにより、前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信することと
を実行し、前記方法は、
前記コンピュータ・プロセッサが、セキュリティの指示に基づいて、前記データの前記複数の要素のエンコーディングをさらに難読化することを決定することと、
前記コンピュータ・プロセッサが、前記データのエンコーディング中に適用するためのデータ・キーを選択することであって、前記データ・キーが、数値を表す２つ以上の要素を含んでいる、前記選択することと、
前記コンピュータ・プロセッサが、エンコーディング・アルゴリズムに関連付けられた１つまたは複数のルールに基づいて、前記データ・キーの少なくとも１つの要素を前記データの各要素に適用することと
をさらに含む、方法。
【請求項４】
前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの前記２つ以上のサブセットに重複せずに分散することが、
前記コンピュータ・プロセッサが、循環的方法で、前記エンコードされたデータの前記複数の要素のその後の要素を、前記エンコードされたデータの前記２つ以上のサブセットに分散することと
をさらに含む、請求項１ないし３の何れか一項に記載の方法。
【請求項５】
データを安全に格納するための方法であって、
コンピュータ・プロセッサが、データを格納するための要求を受信することであって、前記データが複数の要素を含んでいる、前記受信することと、

前記コンピュータ・プロセッサが、１つまたは複数のエンコーディング・アルゴリズムを前記データに適用することによって、エンコードされたデータの複数の要素を生成することであって、前記エンコードされたデータの前記複数の要素の数が、前記データの前記複数の要素の数に等しい、前記生成することと、
前記コンピュータ・プロセッサが、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散することと、
前記コンピュータ・プロセッサが、通信ユニットにより、前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信することと
を実行し、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの前記２つ以上のサブセットに重複せずに分散することが、
前記コンピュータ・プロセッサが、少なくとも第１のエンコーディング・アルゴリズムの構造に基づいて、前記エンコードされたデータの前の要素に対応する値が、前記データのその後の要素をエンコードするために利用されるということを決定することと、
前記コンピュータ・プロセッサが、前記エンコードされたデータの前記その後の要素が、前記エンコードされたデータの前記その後の要素をエンコードするために利用される前記エンコードされたデータの前記前の要素を含んでいる前記エンコードされたデータの別のサブセットと異なるエンコードされたデータのサブセットに割り当てられるように、前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの前記２つ以上のサブセットに分散することと
をさらに含む、方法。

【請求項６】
前記２つ以上のストレージ・システムが、異なる地理的位置でのストレージ・システムである、請求項１ないし５の何れか一項に記載の方法。

【請求項７】
データを安全に格納するためのコンピュータ・プログラムであって、
コンピュータ・プロセッサに、
データを格納するための要求を受信するためのプログラム命令であって、前記データが複数の要素を含んでいる、前記プログラム命令と、
前記データの前記複数の要素を各数値に変換するためのプログラム命令と、
１つまたは複数のエンコーディング・アルゴリズムを、変換された前記データに適用することによって、エンコードされたデータの複数の要素を生成するためのプログラム命令であって、前記エンコードされたデータの前記複数の要素の数が、変換された前記データの前記複数の要素の数に等しい、前記プログラム命令と、
前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散するためのプログラム命令と、
前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信するためのプログラム命令と、を実行させるためのコンピュータ・プログラム。

【請求項８】
データを安全に格納するためのコンピュータ・システムであって、前記コンピュータ・システムが、
１つまたは複数のコンピュータ・プロセッサと、
１つまたは複数のコンピュータ可読ストレージ媒体と、
前記１つまたは複数のコンピュータ・プロセッサのうちの少なくとも１つによって実行するために前記コンピュータ可読ストレージ媒体に格納されたプログラム命令とを備えており、前記プログラム命令が、
データを格納するための要求を受信するためのプログラム命令であって、前記データが複数の要素を含んでいる、前記プログラム命令と、
前記データの前記複数の要素を各数値に変換するためのプログラム命令と、

１つまたは複数のエンコーディング・アルゴリズムを、変換された前記データに適用することによって、エンコードされたデータの複数の要素を生成するためのプログラム命令であって、前記エンコードされたデータの前記複数の要素の数が、変換された前記データの前記複数の要素の数に等しい、前記プログラム命令と、

前記エンコードされたデータの前記複数の要素を前記エンコードされたデータの２つ以上のサブセットに重複せずに分散するためのプログラム命令と、

前記エンコードされたデータの前記２つ以上のサブセットを、対応する数の２つ以上のストレージ・システムに送信するためのプログラム命令とを含む、コンピュータ・システム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、一般に、データ・セキュリティの分野に関連しており、より詳細には、データをエンコードし、エンコードされたデータの一部を異なるクラウドストレージ位置に分散することによってデータを難読化することに関連している。

【背景技術】

【０００２】

クラウド・コンピューティングは、インターネットを経由して、構成可能なコンピューティング・システムのリソースの共有プールへの遍在するアクセスを可能にする情報技術パラダイムである。クラウド・コンピューティング・サービスの一部の利用者（例えば、会社、組織、政府機関）は、さまざまなデータ処理システムを所有していることがあるが、追加の計算リソース（例えば、柔軟な拡張性）を提供するクラウド・コンピューティング・サービスを利用して、コストを削減することができる。クラウド・コンピューティング・サービスのその他の利用者は、ユーザと会社の間のインターフェイスとしてクラウドを利用して、ユーザのより近くでアプリケーションおよび機能を配布することによって応答時間を改善し、唯一のデータ処理の位置で発生する可能性のあるネットワークの混雑を減らし、顧客満足体験を改善する。さまざまなビジネス機能を処理することに加えて、クラウド・コンピューティング・プロバイダは、ＩａａＳ（infrastructure as a service）モデルを介して、個人、インターネットベースのサービス・プロバイダ、および企業に、大容量のクラウドストレージ・ソリューションを提供することができる。１つの例では、企業は、Ｗｅｂサイトを社内でホストすることができるが、ストレージをＩａａＳプロバイダに外部委託することができる。

【０００３】

クラウド・ストレージは、デジタル・データが論理的プールに格納されるコンピュータ・データ・ストレージのモデルである。物理ストレージは、複数のサーバ（場合によっては、複数の位置）に広がることができ、物理的環境は、通常、ホストしている企業（すなわち、サービス・プロバイダ）によって所有され、管理される。加えて、クラウド・コンピューティング・プロバイダは、認証およびアクセスの監視、ファイアウォール、ユーザの分離、侵入の検出および防止、ならびにセキュリティ情報およびイベント・ロギング・システムなどの、さまざまなセキュリティ対策を利用して、格納された情報およびデータを保護する。したがって、クラウドストレージ・プロバイダは、データ・ストレージに関する地理的多様性およびプロバイダの多様性を求めているデータ所有者に、大量のストレージ容量、容量のジャストインタイムの管理、セキュリティ、および耐久性、ならびにデータへの「どこでも、いつでも」可能なアクセスを提供する。

【発明の概要】

【０００４】

本発明の態様は、データを安全に格納するための方法、コンピュータ・プログラム製品、およびシステムを開示する。この方法は、少なくとも１つのコンピュータ・プロセッサが、データを格納するための要求を受信することを含み、このデータは複数の要素を含んでいる。この方法は、少なくとも１つのコンピュータ・プロセッサが、１つまたは複数の

エンコーディング・アルゴリズムをデータに適用することによって、エンコードされたデータの複数の要素を生成することをさらに含み、エンコードされたデータの複数の要素の量は、データの複数の要素の量に等しい。この方法は、少なくとも１つのコンピュータ・プロセッサが、エンコードされたデータの複数の要素を、エンコードされたデータの２つ以上のサブセットに重複せずに分散することをさらに含む。この方法は、少なくとも１つのコンピュータ・プロセッサが、エンコードされたデータの２つ以上のサブセットを、対応する量の２つ以上のストレージ・システムに送信することをさらに含む。

【図面の簡単な説明】

【０００５】

【図１】本発明の一実施形態に従う、クラウド・コンピューティング環境の一部の機能ブロック図である。

【図２】本発明の一実施形態に従ってセキュア・データ分散プログラムの動作ステップを示すフローチャートである。

【図３】本発明の一実施形態に従ってセキュア・データ再構築プログラムの動作ステップを示すフローチャートである。

【図４】本発明の少なくとも１つの実施形態に従って、データ・キーを利用してデータをエンコードし、エンコードされたデータを格納する前にエンコードされたデータをチャンクに分散し、メタデータを利用して、情報をエンコードされたデータのチャンクに関連付ける例を示す図である。

【図５】本発明の一実施形態に従う、コンピュータのコンポーネントのブロック図である。

【図６】本発明の一実施形態に従ってクラウド・コンピューティング環境を示す図である。

【図７】本発明の一実施形態に従ってクラウド・コンピューティング環境の機能的抽象レイヤのセットを示す図である。

【発明を実施するための形態】

【０００６】

本発明の実施形態は、クラウド・コンピューティングによって、企業が情報技術（ＩＴ：information technology）インフラストラクチャの初期費用を防ぐか、または最小限に抑えることができるということを認識する。加えて、本発明の実施形態は、多くの企業が、アプリケーションをクラウドに移行しており、さまざまな企業が、複数のクラウド・プラットフォームおよび複数のサービス・プロバイダをデータおよびアプリケーションに利用しているということを認識する。しかし、データの所有者は、特定のサービス・プロバイダに縛られることを望まないことがある。データの所有者は、データの移行においてデータを失うことなく、著しい努力をせずに、別のプロバイダに切り替えるための柔軟性を必要とする。格納されているデータの重要度が増すことに応じて、格納されているデータの可用性および耐久性を保証するために、データ所有者の期待は、より信頼できるメカニズムが準備されることを必要とするように発展している。本発明の実施形態は、クラウドサービス・プロバイダが、公開された政府統計、地図作成情報、社会的に共有された情報、気象データ、ニュースなどのパブリック・データを、クラウド上にすでに格納しているということを認識する。しかし、本発明の実施形態は、セキュリティの問題、信用の問題、プライバシーに対する懸念、規制などのために、プライベート・データが通常はパブリッククラウドに格納されないということも認識する。このような状況において、データ所有者は、セキュリティ、冗長性をさらに保証し、データ漏洩の脅威の可能性を減らすために、データを単一のクラウド内に格納する代わりに、複数のクラウド・プラットフォームにわたってデータを格納しようとする。

【０００７】

本発明の実施形態は、クラウドストレージ・システム、パブリック・クラウド・システム、またはハイブリッド・クラウド・システム、あるいはその組み合わせなどの、複数のクラウド・プラットフォーム間で分散されるセキュアな形式に基づいて、データを信頼できる形態で安全に格納するための解決策を提供する。本明細書に記載されたデータのセキュアな分散された形式は、チャンクと呼ばれ、さらに具体的には、エンコードされたデー

タのチャンクと呼ばれる。エンコードされたデータの各チャンクは、データのエンコードされた要素のより大きいセットから一意に分散されたデータの要素のサブセット（例えば、より小さいグループ）である。本発明の実施形態は、エンコードされたデータの要素をエンコードされたデータの２つ以上のチャンクに分散する前に、セキュアな分散された格納のためのデータをエンコードする。本発明の実施形態は、１つまたは複数のデータ内の依存関係、再帰的アクション、または元のデータの少なくとも１つの要素もしくはエンコードされるデータの少なくとも１つの前の要素またはその両方の漸進的（すなわち、インデックス付けされた）利用、あるいはその組み合わせを利用して、データの要素をエンコードする。本発明の一部の実施形態は、元のデータをさらに難読化するために、１つまたは複数のエンコード・アルゴリズムによって利用される他の情報から成るデータ・キーも利用する。本発明の実施形態は、可逆的機能を利用して、データをエンコードおよびデコードする。

【０００８】

本発明の実施形態は、エンコードされたデータの単一のチャンクが失われたか、破損したか、または取り出されなかった場合に、元のデータの全体を再構成（例えば、再現、再生）してデコードすることができなくなるように、エンコードされたデータをエンコードされたデータのチャンクに分散することによって、データをさらに保護することができる。追加の実施形態では、エンコードされたデータの個別のチャンクを、分離してデコードすることはできない。本発明のさまざまな実施形態は、大局処理とは対照的に、１つまたは複数のデータ内の要素の依存関係および再帰的動作（例えば、アクション、計算）を利用して、エンコードされたデータのすべてのチャンクが取り出されない限り、元のデータ全体を再構成することを不可能にし、エンコーディング・アルゴリズムおよびデコーディング・アルゴリズムに加えて、エンコードされたデータの分散アルゴリズムが知られている。

【０００９】

したがって、本発明の一部の実施形態は、エンコードされたデータの単一のバイトまたは要素の損失または破損にも応答して、元のデータ全体を再構成することを不可能にする。データ内の要素の依存関係の例は、データの現在の要素の値をエンコードすることが、データの要素の現在の値および少なくとも１つの前のエンコードされたデータの要素の値に基づくことである。現在のデータの値をエンコードするために少なくとも１つの前の元のデータの値を含めることなどの、追加の依存関係はそれぞれ、元のデータをエンコードしてエンコードされたデータのチャンクに分散することに関連する本発明のセキュリティの特徴をさらに強化する。エンコードされたデータをデコードすることも、エンコードされたデータの要素の現在の値および少なくとも１つの前のエンコードされたデータの要素の値などの、データ内の要素の依存関係を含む。少なくとも１つの前のデコードされたデータの要素の値を利用して、エンコードされたデータの現在の要素の値をデコードすることなどの、追加の依存関係はそれぞれ、デコーディング・アルゴリズムのセキュリティをさらに強化する。

【００１０】

本発明のさまざまな実施形態は、エンコードされたデータの１つのチャンクを、元のデータに対応するエンコードされたデータのチャンクを格納するために選択された異なる地理的位置またはクラウドストレージ・プロバイダあるいはその両方の各々に分散することによって、プライベート・データに関連するセキュリティをさらに改善する。１つの例では、少なくとも２つのクラウドストレージ・プロバイダが、同じ都市内で使用できるストレージ容量を有している場合、エンコードされたデータの２つのチャンクが、この都市内で格納されてよい。

【００１１】

本発明の一部の実施形態は、１つまたは複数のクラウドストレージ・プロバイダがデータのバックアップを提供しないなどの場合の災害復旧のために、または１つの地理的位置が問題（例えば、自然災害、停電、ネットワーク崩壊など）による影響を受ける場合でも

、二次ストレージ位置にアクセスすることに基づいて、タイムリーな方法で元のデータがまだ回復可能であることを保証するために、エンコードされたデータのチャンクの重複を利用する。本発明の実施形態は、チェックサムまたはハッシュ関数などの、当該技術において知られた１つまたは複数のデータ検証を利用して、エンコードされたデータまたはデコードされたデータあるいはその両方の取り出されたチャンクの完全性を保証する。

【００１２】

本発明の実施形態の実装は、さまざまな形態を取ってよく、以下では、各図を参照して例示的な実装の詳細が説明される。

【００１３】

ここで、各図を参照して本発明が詳細に説明される。図１は、本発明の１つの実施形態に従って、クラウド・コンピューティング・ノード１０（図６に関してさらに詳細に説明される）などのクラウド・コンピューティング環境（概して、環境１００と指定される）の一部を示す機能ブロック図である。環境１００の実施形態は、ネットワーク１３０を経由してすべて相互接続されたシステム１０２および複数のクラウドストレージ位置（例えば、ストレージ１２０Ａ～ストレージ１２０Ｎ）を含んでいる。図１は、単に１つの実装の例を提供しており、さまざまな実施形態を実装できる環境に関して、どのような制限も意味していない。特許請求の範囲に列挙されている本発明の範囲から逸脱することなく、当業者によって、示された環境に対する多くの変更が行われてよい。

【００１４】

他の実施形態では、図１は、クラウド・コンピューティング環境（概して、環境１００と指定される）とは異なるネットワーク・コンピューティング環境またはストレージ・アーキテクチャの一部を示している機能ブロック図である。１つの例では、ストレージ１２０Ａ～ストレージ１２０Ｎのうちの１つまたは複数は、ネットワーク接続ストレージ（ＮＡＳ：network-attached storage）システムもしくはデバイス、ストレージ・エリア・ネットワーク（ＳＡＮ：storage area networks）、またはその他のストレージ・デバイス、あるいはその組み合わせなどの、さまざまな種類のストレージ・システムを表してよい。

【００１５】

システム１０２およびストレージ１２０Ａ～ストレージ１２０Ｎは、ラップトップ・コンピュータ、タブレット・コンピュータ、ネットブック・コンピュータ、パーソナル・コンピュータ（ＰＣ：personal computers）、デスクトップ・コンピュータ、パーソナル・デジタル・アシスタント（ＰＤＡ：personal digital assistants）、スマートフォン、ウェアラブル・デバイス（例えば、スマート・グラス、スマート・ウォッチなど）、または当該技術において知られた任意のプログラム可能なコンピュータ・システムであってよい。特定の実施形態では、システム１０２およびストレージ１２０Ａ～ストレージ１２０Ｎは、データ・センターおよびクラウド・コンピューティング・アプリケーションにおいて一般的であるように、ネットワーク１３０を介してアクセスされたときにシームレスなリソースの単一のプールとして機能するクラスタ化されたコンピュータおよびコンポーネント（例えば、データベース・サーバ・コンピュータ、アプリケーション・サーバ・コンピュータなど）を利用する、コンピュータ・システムを表す。システム１０２およびストレージ１２０Ａ～ストレージ１２０Ｎは、本発明の実施形態に従って、図５に関してさらに詳細に示され、説明されるように、コンポーネントを含んでよい。一般に、ストレージ１２０Ａ～ストレージ１２０Ｎは、機械可読のプログラム命令を実行し、ネットワーク１３０を介してシステム１０２と通信することができる、任意のプログラム可能な電子デバイスまたはプログラム可能な電子デバイスの組み合わせを表す。

【００１６】

システム１０２は、ストレージ１０３、セキュア・データ分散プログラム２００、およびセキュア・データ再構成プログラム３００を含んでいる。システム１０２は、Ｗｅｂサイト・インターフェイス、１つまたは複数のＷｅｂデーモン、データベース管理システム、ネットワーク・マッピング・プログラム、ネットワーク通信プログラム、さまざまな数

学的変換、統計的変換などを実行するためのさまざまな機能またはプログラムあるいはその両方などの、さまざまなプログラムも含んでいる。１つの実施形態では、システム１０２は、クラウド・コンピューティング・システム内の１つまたは複数のコンピューティング・システムを表す。別の実施形態では、システム１０２は、会社、大学、企業などの組織の１つまたは複数のコンピューティング・システムを表す。一部の実施形態では、システム１０２は、ネットワーク１３０および１つまたは複数の通信プログラムもしくは通信プロトコルまたはその両方を利用して、ストレージ１２０Ａ～ストレージ１２０Ｎをポーリングするか、または照会して、物理的位置、ストレージ・プロバイダの情報、ストレージのコスト、サービス水準合意（ＳＬＡ：service level agreements）、データ・バックアップ機能などの、ストレージ位置の可用性に関連する情報を取得する。システム１０２は、その後、セキュア・データ分散プログラム２００によってアクセスするために、ストレージ情報１０７内に情報を格納してよい。

【００１７】

１つの実施形態では、ストレージ１０３は、データ１０４、データ・キー１０５、セキュリティ・アルゴリズム１０６、およびストレージ情報１０７を含んでいる。ストレージ１０３は、さまざまなテーブル、リスト、連想配列、データベース、またはその他のプログラム、あるいはその組み合わせを含んでもよい（図示されていない）。データ１０４は、分散された格納のためのデータをエンコードすること、およびエンコードされたデータをエンコードされたデータのチャンクに分散することの両方による難読化に基づいて、データのセキュアな分散された格納のためにシステム１０２を利用する１つまたは複数の実体（例えば、ユーザ、プログラム、企業、組織など）に関連付けられた複数のデータを含む。さまざまな実施形態では、データ１０４内のデータは、ネットワーク１３０を介して、さまざまな実体または実行中のプログラム（図示されていない）あるいはその両方によってシステム１０２にアップロードされるか、または転送される。別の実施形態では、データ１０４内のデータは、１つまたは複数の実行中のプログラム（図示されていない）によってシステム１０２内で生成される。

【００１８】

データ・キー１０５は、分散して格納されるデータをエンコードおよびデコードするために１つのセキュリティ・アルゴリズムによって利用される複数のキーを含む。一実施形態では、データ・キー１０５は、さまざまな構造または形式あるいはその両方のデータ・キーを含む。１つの例では、データ・キー１０５内のデータ・キーは、ランダムな数値、英字、記号、またはこれらの組み合わせに基づいてよい。別の例では、データ・キーは、システム１０２にアクセスしているユーザまたはプログラムによって入力されたパスワードまたはパス・フレーズであってよい。一部の実施形態では、データ・キー１０５内のデータ・キーは、π、√２、または分数の最初の１０桁などの、ルールに基づく。一部の実施形態では、分散プログラム２００は、１つまたは複数のプログラム（例えば、関数）を実行して、データ・キー１０５のデータ・キー内の情報をインデックス付けされた数値に変換する。１つのシナリオでは、変換されたデータ・キーが、データ・キー１０５内に格納され、元のデータ・キーにそれぞれ関連付けられる。別のシナリオでは、変換されたデータ・キーが、ストレージ情報１０７内に格納され、エンコードされたデータの複数のチャンクに関連付けられたストレージ位置ＩＤ、チャンクＩＤ、シーケンス・データなどの、分散された格納のためのデータに関連する他の情報に関連付けられる。

【００１９】

セキュリティ・アルゴリズム１０６は、セキュアな分散された格納のためのデータを準備するために利用される複数のアルゴリズムを含む。セキュリティ・アルゴリズム１０６は、データをエンコードし、エンコードされたデータをエンコードされたデータの複数のチャンクに分散し、エンコードされたデータの複数の取り出されたチャンクからエンコードされたデータを再構成（すなわち、再構築）し、エンコードされたデータをデコードするためのアルゴリズムを含む。以下では、本発明を実践するためのアルゴリズムのセットの例が提示され、図４に関してさらに詳細に説明される。セキュリティ・アルゴリズム１

０６内のセキュリティ・アルゴリズムは、システム１０２の管理者によって開発されるか、またはシステム１０２のユーザによって開発されるか、あるいはその両方によって開発され、ユーザに関連付けられたデータに関連して利用されてよい。

【００２０】

さまざまな実施形態では、セキュリティ・アルゴリズム１０６内のデータ・エンコード・アルゴリズムは、ソース・データの構成を利用してさまざまな変換を実行する（例えば、エンコードする）、可逆的機能（すなわち、アルゴリズム）である。１つの実施形態では、セキュリティ・アルゴリズム１０６内の１つまたは複数のアルゴリズムが、データ・キー１０５内のデータ・キーを利用して、データのエンコーディング中にデータの要素をさらに難読化する。一実施形態では、セキュリティ・アルゴリズム１０６は、データ・キー内の情報をインデックス付けされた数値に変換するため、またはデータ・キーを作成するため、あるいはその両方のための１つまたは複数のプログラム（例えば、関数）も含む。

【００２１】

一部の実施形態では、アルゴリズムは１つまたは複数の制約を含む。１つのシナリオでは、データ・エンコード・アルゴリズムは、格納するデータ内のデータのすべての要素の最大値より少なくとも１大きい値でなければならない変数（例えば、ｍａｘ＿ｖａｌ）などの、制約を含む。１つの例では、バイナリ・データを格納することに関して、データの要素のその値の範囲が０～２５５になる可能性がある場合、ｍａｘ＿ｖａｌ＝２５６である。別のシナリオでは、データ・エンコード・アルゴリズムがデータを再帰的にエンコードし、現在のエンコード値が、データの１つまたは複数の前にエンコードされた要素に対応する値に基づく場合（例えば、データ内の要素の依存関係）、データ分散アルゴリズムは、エンコードされたデータの異なるチャンクにおいて相互作用するエンコードされたデータの要素を格納するように制約される。一実施形態では、セキュリティ・アルゴリズム１０６は、データの要素がエンコードされるかどうか、およびどのようにエンコードされるかに影響を与える１つまたは複数の他の基準（例えば、ルール）を含む。１つの例では、エンコード・アルゴリズムに含まれる基準は、Ｋ［ｉ］＝０である場合にＥ［ｉ］＝Ｄ［ｉ］であるということを提示してよい。

【００２２】

セキュリティ・アルゴリズム１０６をさらに参照すると、以下では、分散された格納のためのデータに含まれているデータの複数の要素をエンコードすることと、データのエンコードされた要素をエンコードされたデータのチャンクに分散することと、分散して格納されたデータにアクセスする必要があることに応答して、エンコードされたデータの複数の格納されたチャンクからエンコード・データを取り出して再構成することと、その後、再構成されたエンコード済みデータをデコードしてデータを再生することとのためのセキュリティ・アルゴリズムのセットの実施形態例が提示される。

【００２３】

以下の例示的なアルゴリズムでは、「％」記号はモジュロ（すなわち、モジュラス）演算を表す。例えば、エンコード・アルゴリズムの対は、Ｅ［０］＝（Ｄ［０］＋Ｋ［０］）％ｍａｘ＿ｖａｌを含む（本明細書では、アルゴリズムＥＤ１と呼ばれる）。ここで、Ｅ［０］は、データの第１の要素（インデックス値［ｉ］＝０）のエンコード値を表す。本明細書において使用されるとき、インデックス位置およびインデックス値は、交換可能である。Ｋ［０］は、セキュアな分散された格納のためのデータに割り当てられたデータ・キー１０５のデータ・キーの第１の値（インデックス値［ｉ］＝０）である。本明細書において利用されるモジュロ演算（すなわち、％）に関して、％記号の左側の要素（すなわち、（Ｄ［０］＋Ｋ［０］））に対応する値は、被除数である。％記号の右側の要素（すなわち、ｍａｘ＿ｖａｌ）に対応する値は、除数であり、Ｅ［０］は、データの第１の要素のモジュロ演算によって生成された余りに等しい。一実施形態では、エンコード・アルゴリズムは、セキュアな分散された格納のためのデータのデータの要素の量に等しい（例えば、同じ）量のエンコードされたデータの要素を生成する。

【００２４】

第１のデータ値の後のデータの要素（例えば、ｉ＞０のインデックス値）に対応するデータ値の場合は、Ｅ［ｉ］＝（Ｅ［ｉ－１］＋Ｄ［ｉ－１］＋Ｄ［ｉ］＋Ｋ［ｉ％ｋｅｙ＿ｓｉｚｅ］）％ｍａｘ＿ｖａｌである（本明細書では、アルゴリズムＥＤ２と呼ばれる）。アルゴリズムＥＤ２は、エンコーディング制約（例えば、データ内の要素の依存関係）または少なくとも１つの再帰的アクションあるいはその両方を含む。現在の要素のエンコード値（データ値Ｄ［ｉ］の対応する要素のＥ［ｉ］）は、データ・キーの要素（例えば、Ｋ［ｉ％ｋｅｙ＿ｓｉｚｅ］）によって変更され、前のエンコード値Ｅ［ｉ－１］および前のデータ値Ｄ［ｉ－１］の両方によってさらに変更される。アルゴリズムＥＤ２内で、Ｋ［ｉ％ｋｅｙ＿ｓｉｚｅ］の項は、ｋｅｙ＿ｓｉｚｅ＜ｄａｔａ＿ｓｉｚｅであることに応答して、０～（ｋｅｙ＿ｓｉｚｅ－１）の範囲のインデックス値を生成することに基づいて、データの要素に適用してエンコードするために、データ・キーからの要素の循環的選択を生成する。Ｄａｔａ＿ｓｉｚｅは、セキュアな分散された格納のためのデータの要素の数（すなわち、量）に対応する値である。

【００２５】

セキュリティ・アルゴリズム１０６をさらに参照し、以下では、例示的なエンコード・アルゴリズムに関連付けられた制約に基づいて、関連するデータ「チャンク化」アルゴリズムが説明される。一実施形態では、分散された格納のためのエンコードされたデータを受信するために選択された位置の数（すなわち、量）に基づいて、生成されるチャンクの量（ｎｂ＿ｃｈｕｎｋｓ）は、２以上、かつデータに使用できるストレージ位置の量以下である。１つの例では、データ・チャンク化アルゴリズムおよびデータ要素分散アルゴリズムのセットは、「チャンク」（すなわち、Ｃ）内に含まれている要素の量に対応するサイズ値を含み、このサイズ値は、エンコードされたデータの一部から成り、ｓｉｚｅ［ｃ］＝（ｄａｔａ＿ｓｉｚｅ＋ｎｂ＿ｃｈｕｎｋｓ－１－ｃ）ｄｉｖ　ｎｂ＿ｃｈｕｎｋｓによって表される（本明細書では、アルゴリズムＣＤ１と呼ばれる）。ここで、ｃはチャンクに対応するインデックス値である。説明されたアルゴリズム内の関数「ｄｉｖ」は、被除数（すなわち、（ｄａｔａ＿ｓｉｚｅ＋ｎｂ＿ｃｈｕｎｋｓ－１－ｃ））の値および除数（すなわち、ｎｂ＿ｃｈｕｎｋｓ）の値の整数除算を表す。

【００２６】

一実施形態では、データのエンコードされた要素をチャンクに分散することは、各チャンクおよび各チャンク内のデータの要素に対応するインデックス値を循環的に生成するアルゴリズムに基づく。それに応じて、生成されたインデックス値に対応するデータのエンコードされた要素が、チャンクに分散されるか、または割り当てられる。１つの例では、Ｃ［ｃ］［ｊ］＝Ｅ［ｃ＋ｊ＊ｎｂ＿ｃｈｕｎｋｓ］（本明細書では、アルゴリズムＣＤ２と呼ばれる）は、エンコードされた情報をチャンクに割り当て、０＜＝ｃ＜ｎｂ＿ｃｈｕｎｋｓかつ０＜＝ｊ＜ｓｉｚｅ［ｃ］であり、ここで「ｃ」はエンコードされたデータのチャンクのインデックス値を表し、「ｊ」はエンコードされたデータのチャンク内のエンコードされたデータの要素に対応するインデックス値を表す。一部の実施形態では、チャンクの「ｃ」のインデックス値は、チャンクに対応する識別子（ＩＤ）にそれぞれ関連付けられる。加えて、一部のシナリオでは、チャンクＩＤは、チャンクへのエンコードされたデータの分散に関連するシーケンス（例えば、１、２、３、．．．）に対応する。他のシナリオでは、エンコードされたデータのチャンクのシーケンスは、連想配列などを介して、チャンクＩＤに間接的に関連付けられ得る。

【００２７】

セキュリティ・アルゴリズムの現在のセットを続行し、一実施形態では、チャンクに分散されたデータの複数の要素からエンコードされたデータを再構成する（例えば、分散アルゴリズムを逆転して再現する）ことに対応するアルゴリズムが、０＜＝ｉ＜データサイズの場合に、Ｅ［ｉ］＝Ｃ［ｉ％ｎｂ＿ｃｈｕｎｋｓ］［ｉ　ｄｉｖ　ｎｂ＿ｃｈｕｎｋｓ］によって表される（本明細書では、アルゴリズムＲＥ１と呼ばれる）。ここで、アルゴリズムＲＥ１の［ｉ％ｎｂ＿ｃｈｕｎｋｓ］の項は、チャンク・インデックスまたは関連するチャンクＩＤあるいはその両方に対応し、［ｉ　ｄｉｖ　ｎｂ＿ｃｈｕｎｋｓ］は、チ

ャンク内の要素の設定されたインデックス値に対応する（「ｊ」に関して前に説明されており、０＜＝ｊ＜ｓｉｚｅ［ｃ］である）。

【００２８】

再構成されたエンコード済みデータに適用される対応するデコーディング・アルゴリズムの対が、データ・セキュリティ・アルゴリズムの例示的なセットを完成させる。ここで、ｍａｘ＿ｖａｌの値は、データをエンコードするために利用されるｍａｘ＿ｖａｌの値と同じであり、ここで、インデックス値は、再構成されたエンコード済みデータの要素に対応する。例えば、再構成されたエンコード済みデータの初期の（すなわち、第１の）要素をデコードするために、Ｄ［０］＝（Ｅ［０］－Ｋ［０］＋ｍａｘ＿ｖａｌ）％ｍａｘ＿ｖａｌが利用される（本明細書では、アルゴリズムＤＥ１と呼ばれる）。さらに、再構成されたエンコード済みデータのその後の要素をデコードするために、０＜ｉ＜ｄａｔａ＿ｓｉｚｅの場合に、Ｄ［ｉ］＝（Ｅ［ｉ］－Ｅ［ｉ－１］－Ｄ［ｉ－１］－Ｋ［ｉ％ｋｅｙ＿ｓｉｚｅ］＋ｎｂ＿ｃｈｕｎｋｓ＊ｍａｘ＿ｖａｌ）％ｍａｘ＿ｖａｌである（本明細書では、アルゴリズムＤＥ２と呼ばれる）。データを保護することに関して前述したように、アルゴリズムＤＥ２は、エンコードされたデータのチャンクが欠落しているか、または破損している場合に元のデータの全体を回復不能にするように少なくとも１つの再帰的アクションまたはデータ内の依存関係を含む。

【００２９】

ストレージ情報１０７は、分散された格納のために準備されているか、またはエンコードされたデータのチャンクとしてストレージ１２０Ａ～ストレージ１２０Ｎにすでに格納されているか、あるいはその両方である、複数のデータに関連する情報を含む。１つの実施形態では、ストレージ情報１０７は、データを格納したか、または分散して格納されたデータにアクセスしたか、あるいはその両方である１つまたは複数の実体、ストレージ位置ＩＤ、チャンクＩＤ、エンコードされたデータの複数のチャンクに関連付けられたシーケンス（例えば、分散アルゴリズムに基づくチャンクＩＤシーケンス）、データのエンコーディング中に利用されるデータ・キー、データの完全性検証／情報などの、分散された格納のためのデータに関連する情報を含む。一部の実施形態では、ストレージ情報１０７内の情報は、１つまたは複数のテーブルまたはデータベースあるいはその両方などのデータ構造内に格納される。さまざまな実施形態では、エンコードされたデータのチャンクに関連するメタデータ（図４に関してさらに詳細に説明される）も、ストレージ情報１０７内に含まれ、分散された格納のために、元のデータまたは元のデータに関連付けられた参照にさらに関連付けられる。

【００３０】

別の実施形態では、ストレージ情報１０７内の一部の情報は、ネットワーク１３０を介して複数のストレージ・プロバイダに照会し、物理的位置、ストレージ・プロバイダ情報、ストレージのコストなどの、ストレージ１２０Ａ～ストレージ１２０Ｎに関連する情報を取得するシステム１０２によって、取得される。一実施形態では、ストレージ位置ＩＤは、ストレージ・プロバイダまたはストレージ・システムの対応する物理的（例えば、地理的）位置あるいはその両方にさらに関連付けられる。一部のシナリオでは、ストレージＩＤはシステム１０２によって割り当てられる。他のシナリオでは、ストレージＩＤは、ストレージ１２０Ａ～ストレージ１２０Ｎにそれぞれ関連付けられたストレージ・プロバイダによってシステム１０２に提供された情報に対応する。

【００３１】

一部の実施形態では、セキュア・データ分散プログラム２００およびセキュア・データ再構成プログラム３００は、アプリの所有者または開発者あるいはその両方が、アプリの複数のバージョンのサポートを可能にするために加入するサービスである。セキュア・データ分散プログラム２００およびセキュア・データ再構成プログラム３００は、セキュリティ８６（図６および図７に関してさらに詳細に説明される）などの、クラウド・コンピューティング環境５０の管理レイヤ８０の機能のうちの１つを介して実装されてよい。代替または追加として、セキュア・データ分散プログラム２００およびセキュア・データ再

構成プログラム３００は、データ・セキュリティ９６（図６および図７に関してさらに詳細に説明される）などの、クラウド・コンピューティング環境５０のワークロード・レイヤ９０の仮想化された特徴のうちの１つを介して実装されてよい。別の実施形態では、セキュア・データ分散プログラム２００およびセキュア・データ再構成プログラム３００は、エンコードされたデータのチャンクをシステム１０２に関連付けられた他の内部ストレージ位置に格納するために、メインフレーム、ネットワーク・コンピューティング環境、または分散コンピューティング環境、あるいはその組み合わせなどの、別のアーキテクチャ内で動作しているパブリック・クラウドストレージ環境と異なるストレージ環境内、またはそのような環境に加えたストレージ環境内、あるいはその両方のストレージ環境内で、実装される。

【００３２】

セキュア・データ分散プログラム２００は、パブリック・クラウドストレージ環境などのクラウド内のセキュアな格納のためのデータを準備するプログラムである。他のストレージ環境内で動作するために、セキュア・データ分散プログラム２００のさまざまな態様が実装され得る。１つの実施形態では、セキュア・データ分散プログラム２００は、１つまたは複数のエンコーディング・アルゴリズムを利用して、分散された格納のためのデータの要素を数値形式でエンコードする。セキュア・データ分散プログラム２００は、データ・キーを利用して、データの要素のエンコーディングをさらに難読化してよい。さまざまな実施形態では、セキュア・データ分散プログラム２００は、１つまたは複数の分散アルゴリズムを利用して、データのエンコードされた要素をエンコードされたデータのチャンクに分散する。一実施形態例では、エンコードされたデータをエンコードされたデータのチャンクに分散することによって、セキュア・データ分散プログラム２００は、元のデータに関連付けられたエンコードされたデータのチャンクのすべてが取り出され、エンコードされたデータのチャンクへのデータのエンコードされた要素の分散順序も知られない限り、元のデータ全体を回復不能にすることができる。

【００３３】

別の実施形態では、セキュア・データ分散プログラム２００が、分散された格納のためのデータがすでに格納されており、データの現在のバージョンが変更されているということを決定した場合、セキュア・データ分散プログラム２００は、データの現在のバージョンに対応するエンコードされたデータのチャンクを格納する前に、バックアップ指示がないことに基づいて、分散されたストレージ内のエンコードされたデータのチャンクを削除する。セキュア・データ分散プログラム２００は、１つまたは複数のストレージ位置を再利用して、データの現在のバージョンに対応するエンコードされたデータのチャンクを格納してよく、またはセキュア・データ分散プログラム２００は、データの現在のバージョンに対応するエンコードされたデータのチャンクを格納するためのストレージの異なるセットを選択してよい。一実施形態では、セキュア・データ分散プログラム２００は、エンコードされたデータのチャンクをストレージ位置のセットに送信する前に、１つまたは複数のセキュリティ・アルゴリズムの性能を改善するために、分散された格納のためのデータをシステム１０２のＲＡＭ（図示されていない）に移行する。

【００３４】

セキュア・データ再構成プログラム３００は、分散して格納されたデータを識別し、分散して格納されたデータを含んでいるエンコードされたデータの複数のチャンクを取り出し、エンコードされたデータの複数の取り出されたチャンクからエンコードされたデータを再構成した後に、エンコードされたデータをデコードするプログラムである。１つの実施形態では、セキュア・データ再構成プログラム３００は、２つ以上のクラウドストレージ位置に分散されたエンコードされたデータの複数のチャンクからデータを再構成するように動作する。一部の実施形態では、セキュア・データ再構成プログラム３００は、ユーザまたはプログラムあるいはその両方が、分散して格納されたデータにアクセスすることを示していることに応答して、セキュリティ・プログラム、関数、またはＡＰＩ、あるいはその組み合わせによって自動的に実行される。

【００３５】
さまざまな実施形態では、セキュア・データ再構成プログラム３００は、エンコードされたデータの複製されたチャンクまたはバックアップされたチャンクを利用して、エンコードされたデータを再構成する。セキュア・データ再構成プログラム３００は、１つまたは複数のデータ完全性検証方法を利用して、エンコードされたデータの複製されたチャンクが、エンコードされたデータの欠落、または破損したチャンクの正確な複製であることを保証することができる。一実施形態では、セキュア・データ再構成プログラム３００は、１つまたは複数のセキュリティ・アルゴリズムの性能を改善するために、エンコードされたデータまたは再構成されたエンコード済みデータあるいはその両方の取り出されたチャンクを、システム１０２のＲＡＭ（図示されていない）内に格納する。

【００３６】
１つの実施形態では、システム１０２は、ネットワーク１３０を介して、ストレージ１２０Ａ～ストレージ１２０Ｎと通信する。別の実施形態では、ネットワーク１３０は、システム１０２を、ソフトウェア契約の監視（図７に関してさらに詳細に説明される）などの１つまたは複数のサービスを含んでいるクラウド・コンピューティング環境５０（図６に関してさらに詳細に説明される）の１つまたは複数のノードに接続してもよい。１つの例では、ネットワーク１３０は、本発明の１つまたは複数の実施形態を利用するためのライセンスが付与されたアプリに関連するネットワーク・トラフィックをシステム１０２にルーティングするネットワーク・エージェント／デーモンを含んでよい。代替として、ネットワーク１３０は、本発明の実施形態を利用するためのライセンスが付与されていないアプリに関連するネットワーク・トラフィックを、クラウド・コンピューティング環境５０内の別のノードにルーティングしてよい。

【００３７】
例えば、ネットワーク１３０は、ローカル・エリア・ネットワーク（ＬＡＮ：local area network）、電気通信ネットワーク、イントラネットなどの無線ローカル・エリア・ネットワーク（ＷＬＡＮ：wireless local area network）、インターネットなどの広域ネットワーク（ＷＡＮ：wide area network）、またはこれらの任意の組み合わせであることができ、有線接続、無線接続、または光ファイバ接続を含むことができる。一般に、ネットワーク１３０は、本発明の実施形態に従って、システム１０２とストレージ１２０Ａ～ストレージ１２０Ｎの間の通信をサポートする接続およびプロトコルの任意の組み合わせであることができる。さまざまな実施形態では、ネットワーク１３０は、有線接続、無線接続、または光学的接続を介してローカルに動作し、接続およびプロトコル（例えば、パーソナル・エリア・ネットワーク（ＰＡＮ：personal area network）、近距離無線通信（ＮＦＣ：near field communication）、レーザー、赤外線、超音波など）の任意の組み合わせであることができる。一部の実施形態では、ネットワーク１３０の実例は、クラウド・コンピューティング環境５０の２つ以上の仮想化された実体（図示されていない）を接続する仮想ＬＡＮ（ＶＬＡＮ：virtual LAN）である。他の実施形態では、ネットワーク１３０は、ネットワーク・コンピューティング環境である環境１００の実例内の１つまたは複数のＬＡＮまたはＶＬＡＮあるいはその両方を表し、システム１０２とストレージ１２０Ａ～ストレージ１２０Ｎの間の通信を可能にする。

【００３８】
１つの実施形態では、ストレージ１２０Ａ～ストレージ１２０Ｎは、それぞれ異なる地理的位置の複数のストレージ・システムを表す。さまざまな実施形態では、ストレージ１２０Ａ～ストレージ１２０Ｎのうちの１つまたは複数は、クラウドストレージ・システム内またはクラウド・コンピューティング環境内あるいはその両方に含まれる。一部の実施形態では、ストレージ１２０の各実例をホストするクラウドストレージ・システムまたはクラウド・コンピューティング環境あるいはその両方は、２つ以上の異なるベンダーまたはクラウドサービス・プロバイダに対応する。別の実施形態では、ストレージ１２０Ａ～ストレージ１２０Ｎは、システム１０２に関連付けられたネットワーク・コンピューティング環境、分散コンピューティング環境などの複数のストレージ・システムまたはストレ

ージ・デバイスあるいはその両方を表す。

【００３９】

図２は、セキュア・データ分散プログラム２００の動作ステップを示すフローチャートであり、このプログラムは、本発明の実施形態に従って、データをエンコードし、エンコードされたデータをデータのより小さいセットに分散し、データのより小さいセットを複数のストレージ位置に送信することによって、データを安全に格納する。さまざまな実施形態では、セキュア・データ分散プログラム２００によって識別されたか、または更新されたか、あるいはその両方を実行された情報の一部が、その後、分散して格納されたデータを取り出して再構成するために、セキュア・データ分散プログラム３００の態様によって利用される。

【００４０】

ステップ２０２で、セキュア・データ分散プログラム２００が、分散された格納のためのデータを識別する。１つの実施形態では、セキュア・データ分散プログラム２００は、ユーザ、プログラム、セキュリティ機能から取得された情報に基づいて、またはデータに関連付けられたセキュリティ・ルールに基づいて、あるいはその両方に基づいて、分散された格納のためのデータをデータ１０４内で識別する。一部の実施形態では、セキュア・データ分散プログラム２００は、データを分散するためのストレージ位置の量に対応する値または値の範囲、データ・キー（例えば、キー長、キーＩＤ、データ・キーに関連するルールなど）に関連付けられた指示、格納するデータを含んでいる要素から識別された最大値などの、セキュアな分散された格納のためのデータに関連付けられた他の情報を決定する。別の実施形態では、セキュア・データ分散プログラム２００は、データ・キー１０５内、セキュリティ・アルゴリズム１０６内、またはストレージ情報１０７内、あるいはその組み合わせ内の、セキュアな分散された格納のためのデータに関連付けられた情報を決定する。セキュア・データ分散プログラム２００は、コストまたは政府規制などの、セキュアな分散された格納のためのデータに関連付けられた制約を識別してよい。

【００４１】

一実施形態では、セキュア・データ分散プログラム２００は、システム１０２にインストールされたさまざまな機能またはプログラム（図示されていない）あるいはその両方を利用して、さまざまな数学的変換、統計的変換などを実行し、データに対応するサイズ、データに関連するｍａｘ＿ｖａｌ、キーおよびデータの要素のインデックス値などの、分散された格納のためのデータに関連する情報を取得し、分散された格納のためのデータに関連付けられた情報をリンクするテーブル／連想配列を生成してよい。さまざまな実施形態では、セキュア・データ分散プログラム２００は、データ・エンコーディング・アルゴリズム、データ・チャンク化アルゴリズム、およびデータ分散アルゴリズムなどの、分散された格納のためのデータに関連して利用されるアルゴリズムのセットも識別する。

【００４２】

ステップ２０４で、セキュア・データ分散プログラム２００は、データの分散されたストレージの位置のセットを決定する。一実施形態では、セキュア・データ分散プログラム２００は、ストレージ情報１０７内の使用できる情報に基づいて、データの分散されたストレージの位置またはストレージＩＤあるいはその両方のセットを決定する。その後、セキュア・データ分散プログラム２００は、ストレージ情報１０７内のデータの分散されたストレージに関連する他の情報を含むか、または更新するか、あるいはその両方を実行してよい。さまざまな実施形態では、セキュア・データ分散プログラム２００は、ネットワーク１３０を利用して、１つまたは複数のクラウドストレージ・プロバイダをポーリングし、分散された格納のためのデータを含んでいるエンコードされたデータのチャンクを格納するために使用できるストレージ１２０Ａ～ストレージ１２０Ｎなどの位置を識別する。１つの実施形態では、セキュア・データ分散プログラム２００は、分散された格納のためのデータ（ステップ２０２）に関連付けられたストレージ位置の量に関連する情報に基づいて、位置のセットを決定する。

【００４３】

別の実施形態では、セキュア・データ分散プログラム２００は、１つまたは複数のクラウドストレージ・プロバイダをポーリングすることに応答して識別された使用可能な位置の量に基づいて、データの分散された格納のための位置のセットを決定する。一部の実施形態では、セキュア・データ分散プログラム２００は、分散された格納のためのデータに関連付けられた情報によって指示されたストレージ位置の量に加えて、１つまたは複数のストレージ位置を識別する。その後、セキュア・データ分散プログラム２００は、１つまたは複数の制約またはルールを利用して、データの分散されたストレージの位置のセットを決定する。他の実施形態では、セキュア・データ分散プログラム２００は、クラウドストレージ環境と異なる１つまたは複数のストレージ・アーキテクチャに基づいて、データの分散されたストレージのストレージ位置のセットを決定する。

【００４４】

ステップ２０６で、セキュア・データ分散プログラム２００が、分散された格納のためのデータをエンコードする。セキュア・データ分散プログラム２００は、１つまたは複数のエンコード・アルゴリズムを利用して、分散された格納のための元のデータの複数の要素にそれぞれ関連付けられた、エンコードされたデータの複数の要素を生成する。１つの実施形態では、セキュア・データ分散プログラム２００は、セキュアな分散された格納のためのデータに関連付けられた情報によって指示されたセキュリティ・アルゴリズム１０６内の１つまたは複数のエンコード・アルゴリズムを利用して、分散された格納のためのデータをエンコードする。別の実施形態では、セキュア・データ分散プログラム２００は、セキュリティ・アルゴリズム１０６から選択された１つまたは複数のエンコード・アルゴリズムを利用して、分散された格納のためのデータをエンコードする。

【００４５】

一部の実施形態では、セキュア・データ分散プログラム２００は、１つまたは複数のエンコード・アルゴリズムを利用して、データ・キーからの１つまたは複数の数値を適用することによって、エンコードされたデータをさらに難読化する（図４に関してさらに詳細に説明される）。セキュア・データ分散プログラム２００が、データ・キーが１つまたは複数のルールに基づいているか、または非数値の文字を含んでいるか、あるいはその両方であるということを決定した場合、セキュア・データ分散プログラム２００は、１つまたは複数の変換またはルールあるいはその両方を利用して、数値を含んでいるデータ・キーのバージョンを生成する。さまざまな実施形態では、セキュア・データ分散プログラム２００は、ｍａｘ＿ｖａｌ変数、データ・キー１０５からのデータ・キー、１つまたは複数の利用されるエンコード・アルゴリズムを識別する情報、データ・キーの数値のバージョン、エンコードされたデータに対応するチェックサムまたはハッシュ値などの、分散された格納のためのデータをエンコードすることに関連する識別された情報に基づいて、セキュアな分散された格納のためのデータに関連付けられた情報を更新する。

【００４６】

ステップ２０８で、セキュア・データ分散プログラム２００は、エンコードされたデータを、エンコードされたデータの複数のチャンクに分散する。一実施形態では、セキュア・データ分散プログラム２００は、エンコードされたデータを、ストレージ位置の決定されたセット内のストレージ位置の量に対応する量（例えば、ｎｂ＿ｃｈｕｎｋｓ）の、エンコードされたデータのチャンクに分散する。セキュア・データ分散プログラム２００は、ＩＤ、シーケンス情報、およびエンコードされたデータのチャンクに対応するサイズ値（図４に関してさらに詳細に説明される）、エンコードされたデータのチャンクごとのチェックサムまたはハッシュ値などの情報を、さらに決定するか、または関連付けるか、あるいはその両方を実行してよい。

【００４７】

一部の実施形態では、セキュア・データ分散プログラム２００は、（セキュリティ・アルゴリズム１０６に関して前に説明された）分散された格納のためのデータに関連付けられた情報によって指示された、セキュリティ・アルゴリズム１０６内の１つまたは複数のデータ・チャンク化アルゴリズムおよびデータ分散アルゴリズムを利用して、エンコード

されたデータを分散する。別の実施形態では、セキュア・データ分散プログラム２００は、セキュリティ・アルゴリズム１０６から選択された１つまたは複数のデータ・チャンク化アルゴリズムおよびデータ分散アルゴリズムを利用して、エンコードされたデータを分散する。さまざまな実施形態では、セキュア・データ分散プログラム２００が、利用されるエンコード・アルゴリズムがデータ内の要素の依存関係を含んでいるということを決定した場合、セキュア・データ分散プログラム２００は、データ内の要素の依存関係と互換性があるデータ分散アルゴリズムを利用する（図４に関してさらに詳細に説明される）。

【００４８】

ステップ２１０で、セキュア・データ分散プログラム２００は、位置のセット内の格納のために、エンコードされたデータの複数のチャンクを送信する。さまざまな実施形態では、セキュア・データ分散プログラム２００は、１つまたは複数の通信プロトコルを実行し、ネットワーク１３０を介して、エンコードされたデータのチャンクを格納するクラウドベースのストレージ位置の（例えば、ストレージ１２０Ａ～ストレージ１２０Ｎから）決定されたセットとの通信を確立し、認証する（例えば、アクセスを確立する）。１つの実施形態では、セキュア・データ分散プログラム２００が、エンコードされたデータの送信されたチャンクの各々が正常に格納されたという確認応答を受信することに応答して、セキュア・データ分散プログラム２００は、元のデータをデータ１０４から削除する。一部の実施形態では、セキュア・データ分散プログラム２００が、エンコードされたデータのチャンクをストレージ位置の決定されたセットからのストレージ位置内に格納できない場合（例えば、認証の失敗、タイムアウト、接続の問題、問題、別の可用性の問題など）、セキュア・データ分散プログラム２００は、決定された１つまたは複数の追加の（例えば、代替の）ストレージ位置から、エンコードされたデータのチャンクの別のストレージ位置を選択する。

【００４９】

１つの実施形態では、セキュア・データ分散プログラム２００が、エンコードされたデータの送信されたチャンクの各々が正常に格納されたという確認応答を受信することに応答して、セキュア・データ分散プログラム２００は、エンコードされたデータのチャンクに対応する位置を識別するストレージ情報１０７を更新する。セキュア・データ分散プログラム２００は、データの分散されたストレージに関連する情報を格納するメタデータ情報またはその他のデータ構造（例えば、メタデータ、テーブル、連想配列など）あるいはその両方を更新するなどの、データの分散されたストレージに関連する情報に基づいて、ストレージ情報１０７を更新する。

【００５０】

図３は、セキュア・データ再構成プログラム３００の動作ステップを示すフローチャートであり、このプログラムは、本発明の実施形態に従って、セキュアな格納のために複数の位置に分散されたエンコードされたデータの複数のチャンクを取り出し、再び組み立て、デコードすることによって、セキュアな格納のために分散されたデータを復元する。

【００５１】

ステップ３０２で、セキュア・データ再構成プログラム３００は、分散して格納されたデータに関連する情報を識別する。セキュア・データ再構成プログラム３００は、ユーザ、ソフトウェア・アプリケーション、またはシステムの機能、あるいはその組み合わせなどの、システム１０２を利用している１つまたは複数の実体から要求を受信することに基づいて、分散してシステム１０２に格納されたデータを取り出して復元するための要求に応答してよい。一実施形態では、セキュア・データ再構成プログラム３００は、ストレージ情報１０７内の情報に基づいて、分散して格納されたデータに関連する情報を識別する。１つの例では、セキュア・データ再構成プログラム３００は、チャンクＩＤ、エンコードされたデータのチャンクに関連するシーケンス（例えば、チャンクＩＤシーケンス）、エンコードされたデータのチャンクに対応するストレージ位置、分散された格納のためのデータを準備するために利用されるセキュリティ・アルゴリズムのセット、データのエンコーディング中に適用されるデータ・キーなどの、分散して格納されたデータに関連する

ストレージ情報１０７内の情報を識別する。セキュア・データ再構成プログラム３００は、再構成されたデータを格納するための指定された位置を識別してもよい。一部の実施形態では、セキュア・データ再構成プログラム３００は、エンコードされたデータのチャンクの追加のコピーを分散するために、セキュア・データ分散プログラム２００がバックアップ（例えば、複製）方針を利用したかどうか、データ完全性検証の情報、およびエンコードされたデータのチャンクの追加のコピーのストレージに関連する情報などの、分散されたデータの他の情報も識別する。

【００５２】

ステップ３０４で、セキュア・データ再構成プログラム３００は、分散して格納されたデータに対応するエンコードされたデータのチャンクを取り出す。セキュア・データ再構成プログラム３００は、ストレージ情報１０７内の分散して格納されたデータに関連する識別された情報を利用して、分散して格納されたデータに対応するエンコードされたデータのチャンクに対応するストレージ位置を識別する。一部の実施形態では、セキュア・データ再構成プログラム３００は、１つまたは複数の通信プロトコルを実行し、ネットワーク１３０を介して、取り出すエンコードされたデータのチャンクを格納するストレージ１２０の実例との通信を確立し、認証する（例えば、アクセスを確立する）。さまざまな実施形態では、セキュア・データ再構成プログラム３００は、ストレージ１２０Ａ～ストレージ１２０Ｎから識別されたクラウドストレージ位置（例えば、異なる地理的位置）から、エンコードされたデータのチャンクを取り出す。

【００５３】

別の実施形態では、セキュア・データ再構成プログラム３００が、分散して格納されたデータに対応するエンコードされたデータのチャンクを正常に取り出すことに応答して、セキュア・データ再構成プログラム３００は、エンコードされたデータのチャンクを各ストレージ位置から削除する。代替として、セキュア・データ再構成プログラム３００は、分散して格納されたデータに関連する情報に基づいて、エンコードされたデータのチャンクを各ストレージ位置内に残してよい。

【００５４】

一実施形態では、セキュア・データ再構成プログラム３００が、分散して格納されたデータに対応するエンコードされたデータのチャンクを格納する１つまたは複数の位置がアクセス可能でないということを決定した場合（例えば、ネットワークの問題、サーバの問題など）、セキュア・データ再構成プログラム３００は、セキュア・データ分散プログラム２００が、複製方針を利用して、分散して格納されたデータに対応するエンコードされたデータのチャンクを分散したかどうかを判定する。セキュア・データ分散プログラム２００が複製方針を利用したということの決定に応答して、セキュア・データ再構成プログラム３００は、一次ストレージ位置から取り出すことができないエンコードされたデータのチャンクを格納している１つまたは複数の他の（例えば、二次的な、代替の）位置を、ストレージ１２０Ａ～ストレージ１２０Ｎから識別する。

【００５５】

ステップ３０４をさらに参照すると、他の実施形態では、セキュア・データ再構成プログラム３００は、さまざまな通信方法を利用して、異なるストレージ・ノード、サーバ、ＳＡＮシステム、またはＮＡＳシステム、あるいはその組み合わせなどの、クラウドストレージ位置と異なるか、またはクラウドストレージ位置に加えて、エンコードされたデータのチャンクを格納しているストレージ１２０Ａ～ストレージ１２０Ｎの１つまたは複数の他の実例へのアクセスを確立する。一実施形態では、ストレージ１２０Ａ～ストレージ１２０Ｎの他の実例のうちの１つまたは複数は、同じ地理的位置または物理的位置あるいはその両方にあることができる。

【００５６】

ステップ３０６で、セキュア・データ再構成プログラム３００は、エンコードされたデータの取り出されたチャンクから、エンコードされたデータを再構成する。一実施形態では、セキュア・データ再構成プログラム３００は、セキュリティ・アルゴリズム１０６内

に含まれている、データを再構成するための１つまたは複数のアルゴリズム、および分散して格納されたデータに関連する情報を利用して、エンコードされたデータのチャンクに分散されたエンコードされたデータの複数の要素から、エンコードされたデータを再構成する。１つの実施形態では、セキュア・データ再構成プログラム３００は、チャンクＩＤまたはチャンク・シーケンス情報あるいはその両方を利用して、分散して格納されたデータに対応するエンコードされたデータを再構成する（例えば、再び組み立てる、再構築する、再生する）。別の実施形態では、セキュア・データ再構成プログラム３００は、エンコードされたデータのチャンクに関連付けられたメタデータを利用して、分散して格納されたデータに対応するエンコードされたデータを再構成する。さまざまな実施形態では、セキュア・データ再構成プログラム３００は、分散して格納されたデータに対応する元のエンコードされたデータに適用されたデータ・チャンク化アルゴリズムおよびデータ分散アルゴリズムのプロセスを逆転する１つまたは複数の再構成アルゴリズムを利用する。

【００５７】

ステップ３０８で、セキュア・データ再構成プログラム３００は、再構成されたエンコード済みデータをデコードする。セキュア・データ再構成プログラム３００は、分散して格納されたデータに対応する再構成されたエンコード済みデータをデコードする。さまざまな実施形態では、セキュア・データ再構成プログラム３００は、分散して格納されたデータに対応するエンコードされたデータをデコードすることに関連する、ステップ３０２で取得された他の情報を識別する。１つの例では、セキュア・データ再構成プログラム３００は、分散された格納のためのデータをエンコードするために利用されるデータ・キー、データ・キーに関連する情報（例えば、キー・サイズ、データ・キー生成ルールなど）、および分散された格納のためのデータをエンコードするために利用されたエンコード・アルゴリズムに対応するデコード・アルゴリズムを識別する。１つの実施形態では、セキュア・データ再構成プログラム３００は、データ・キー１０５内に格納されている識別されたデータ・キーおよびセキュリティ・アルゴリズム１０６内に含まれている対応するデコード・アルゴリズムを利用することによって、再構成されたエンコード済みデータをデコードする。別の実施形態では、セキュア・データ再構成プログラム３００は、ストレージ情報１０７に含まれ、分散して格納されたデータに関連付けられている、識別されたデータ・キーまたはデータ・キー生成ルールあるいはその両方を利用することによって、再構成されたエンコード済みデータをデコードする。

【００５８】

図４は、本発明の実施形態に従って、セキュリティ・アルゴリズムのセット（例えば、アルゴリズムＥＤ１、ＥＤ２、ＣＤ１、ＣＤ２、ＲＥ１、ＤＥ１、およびＤＥ２）を利用して、データの複数の要素をエンコードし、その後、エンコードされたデータのチャンクのエンコードされた要素を分散することに基づいて、セキュアな分散された格納のためのデータの複数の要素を含んでいるデータを準備する例である。図４は、セキュアな分散された格納のためのデータに適用されるセキュリティ・アルゴリズムのさまざまな態様の一部の出力も示している。

【００５９】

データ４０１は、セキュアな分散された格納のためのデータの例である。一実施形態では、データ４０１は、「Ａ」などの、名前またはその他の識別子（例えば、データＩＤ）に関連するメタデータ（図示されていない）を関連付けている。データ４０１に関連付けられたメタデータは、ストレージ情報１０７に含まれてよい。この例では、データ４０１は、データの１５個の要素から成る。この例では、行４０２は、データ４０１のデータの要素に対応する０～１４の範囲のインデックス位置（例えば、インデックス値）を示している。図に示された例では、インデックス位置ｉ＝０で、データ４０１内のデータの対応する要素の値が１４であり、ｉ＝１０で、データの対応する要素の値が１である。一実施形態では、セキュア・データ分散プログラム２００が、データ４０１のデータをエンコードするために、（ステップ２０２で）ｍａｘ＿ｖａｌ＝２４（例えば、２３＋１）を決定する。加えて、セキュア・データ分散プログラム２００が、データ４０１内のデータの要

素の量に対応する値（すなわち、ｄａｔａ＿ｓｉｚｅ＝１５）を決定する。

【００６０】

加えて、セキュア・データ分散プログラム２００は、データ・キーに関連付けられた指示などの、データ４０１に関連付けられた情報に基づいて、データ４０１に割り当てられたエンコード・アルゴリズムがｋｅｙ＿ｓｉｚｅ＝３のキーを指示するということも決定する。図に示された例では、キー４０３は、０、１、および２の各インデックス位置を有する３つの要素３、１、および９から成る。一実施形態では、キー４０３の３つの要素が数値であるため、セキュア・データ分散プログラム２００は、キー４０３内の情報の追加の変換も操作も必要ではないということを決定する。

【００６１】

一実施形態では、セキュア・データ分散プログラム２００は、（セキュリティ・アルゴリズム１０６に関して前に説明された）アルゴリズムＥＤ１およびアルゴリズムＥＤ２（Ｅ［０］＝（Ｄ［０］＋Ｋ［０］）％ｍａｘ＿ｖａｌおよびＥ［ｉ］＝（Ｅ［ｉ－１］＋Ｄ［ｉ－１］＋Ｄ［ｉ］＋Ｋ［ｉ％ｋｅｙ＿ｓｉｚｅ］）％ｍａｘ＿ｖａｌ）を利用して、セキュアな分散された格納のためのデータ４０１をエンコードするということを（ステップ２０２で）決定する。

【００６２】

１つの例では、セキュア・データ分散プログラム２００は、データ４０１を（ステップ２０６で）エンコードする。一実施形態では、例示的なエンコード関数のＫ［ｉ％ｋｅｙ＿ｓｉｚｅ］の項は、データ・キー４０３からの値をデータ４０１の要素に循環的に適用してエンコードする。この例では、列４０４が、キー４０３に適用されたＫ［ｉ％ｋｅｙ＿ｓｉｚｅ］の項の出力である。例示的なエンコード・アルゴリズムの対内のＫ［０］およびＫ［ｉ］の項に基づいて、列４０４の各要素が、対応するインデックス値を有するデータ４０１の要素のデータ値にそれぞれ適用される。例えば、ｉ＝６の場合、Ｋ［６］＝３であり、この値が、データ４０１のデータＤ［６］の要素（すなわち、１７）をエンコードすることに応答して利用される。

【００６３】

この例では、エンコードされたデータ４１０は、キー４０３ならびにアルゴリズムＥＤ１およびアルゴリズムＥＤ２を利用してデータ４０１をエンコードするセキュア・データ分散プログラム２００の出力を示している。一実施形態では、データ４０１のデータの要素のインデックス値は、エンコードされたデータ４１０のデータの要素のインデックス値にも対応している。一部の実施形態では、セキュア・データ分散プログラム２００は、指示されたエンコード・アルゴリズムのうちの少なくとも１つがデータ内の要素の依存関係を含むということも（ステップ２０６で）決定する。例えば、上記のエンコード・アルゴリズムは、前のエンコード値（すなわち、Ｅ［ｉ－１］）およびデータの要素の対応する前の値（すなわち、Ｄ［ｉ－１］）に関連付けられたデータ内の要素の依存関係を含んでいる。

【００６４】

一実施形態では、セキュア・データ分散プログラム２００は、分散されたデータ（すなわち、データ４０１）を受信して格納するためのサイトの量に対応する値が４であり、したがって、作成するデータまたはチャンクのサブセットの量が４である（例えば、ｎｂ＿ｃｈｕｎｋｓ＝４）ということを（ステップ２０４で）決定する。さらに、セキュア・データ分散プログラム２００が、データ４０１をエンコードするために利用されるアルゴリズム（すなわち、アルゴリズムＥＤ２）が、１つまたは複数のデータ内の要素の依存関係を含んでいるということを決定することに応答して、セキュア・データ分散プログラム２００は、エンコードされたデータ４１０が４つのチャンクに分散されるということ、およびデータの各チャンクが、エンコードされたデータ４１０のＥ［５］およびＥ［６］などの連続的な対応するインデックス値を有するデータ要素を含むことを制限されるということを決定する。一部の実施形態では、チャンクの「ｃ」のインデックス値は、要素４２０のチャンクＩＤなどの、分散された格納のためのデータのチャンクに対応する識別子（Ｉ

Ｄ）に、それぞれ関連付けられる。

【００６５】

この例では、チャンクＩＤ１がチャンク４３０Ａに関連付けられ、チャンクＩＤ２がチャンク４３０Ｂに関連付けられ、チャンクＩＤ３がチャンク４３０Ｃに関連付けられ、チャンクＩＤ４がチャンク４３０Ｄに関連付けられる。一部の実施形態では、セキュア・データ分散プログラム２００は、データ４０１の分配された格納に関連付けられたセキュリティ・アルゴリズムのセットが、１つまたは複数のデータ・チャンク化アルゴリズムを含むということを識別する。データ・チャンク化アルゴリズムの１つの態様は、チャンクに対応するインデックス値に基づいて、「チャンクＩＤ」内に含まれている要素の量に対応するサイズ値を決定する。一実施形態例では、チャンクに対応するサイズが、アルゴリズムＣＤ１によって決定される。１つの例では、チャンク４３０Ｂに関連付けられたチャンクＩＤ２は、チャンク・インデックス値（chunk index value）＝１と、エンコードされたデータ４１０の４つの要素がチャンク４３０Ｂに割り当てられることを示す対応するサイズ値＝４とを有する。

【００６６】

一実施形態では、セキュア・データ分散プログラム２００は、エンコードされたデータ４１０を、エンコードされたデータの４つのチャンクに循環的に分散する。図に示された例では、セキュア・データ分散プログラム２００は、データ・チャンク化アルゴリズムを利用して、データの４つのチャンクへの分散のために、エンコードされたデータ４１０内のエンコードされたデータの要素に対応するインデックス値を循環的に決定する。それに応じて、セキュア・データ分散プログラム２００は、決定されたインデックス値に対応するエンコードされたデータの要素を、チャンクＩＤに対応するチャンクに（ステップ２０８で）分散する。この例では、セキュア・データ分散プログラム２００は、アルゴリズムＣＤ２を利用して、エンコードされたデータ４１０を、１、２、３、および４の対応するＩＤを有するチャンクに分散し、ここで、「ｃ」はチャンクのインデックス値に対応し、「ｊ」はチャンク内の要素に対応するインデックス値を表す。例えば、データの第３のチャンク（すなわち、チャンク４３０Ｃ）の最後の要素のインデックス値をアルゴリズムＣＤ２に入力すると、Ｃ［２］［３］＝Ｅ［２＋３＊４］＝Ｅ［１４］が得られ、この値は、エンコードされたデータ４１０内の２１の値に等しい。

【００６７】

この例では、要素４２０に関連付けられたチャンクＩＤに対応する網掛けによって、循環的分散データのエンコードされたデータ４１０が示されている。一部の実施形態では、セキュア・データ分散プログラム２００は、データ４０１に関連付けられたエンコードされたデータのチャンクの順序を示すチャンクＩＤを利用する。１つの例では、網掛けなしで示されたエンコードされたデータ４１０の要素がチャンクＩＤ１に対応し（すなわち、１７、１５、２０、および２のエンコード値）、斑点の網掛けで示されたエンコードされたデータ４１０の要素がチャンクＩＤ２に対応し（すなわち、１０、７、２２、および１９のエンコード値）、クロスハッチングの網掛けで示されたエンコードされたデータ４１０の要素がチャンクＩＤ３に対応し（すなわち、２２、１２、０、および２１のエンコード値）、高密度の斑点の網掛けで示されたエンコードされたデータ４１０の要素がチャンクＩＤ４に対応する（すなわち、９、９、および１４のエンコード値）。

【００６８】

この例では、チャンク４３０Ａは、エンコードされたデータの要素のセットと、チャンクＩＤ１のエンコードされたデータ４１０の要素の対応する順序とを表している。同様に、チャンク４３０Ｂは、エンコードされたデータの要素のセットと、チャンクＩＤ２のエンコードされたデータ４１０の要素の対応する順序とを表しており、チャンク４３０Ｃは、エンコードされたデータの要素のセットと、チャンクＩＤ３のエンコードされたデータ４１０の要素の対応する順序とを表しており、チャンク４３０Ｄは、エンコードされたデータの要素のセットと、チャンクＩＤ４のエンコードされたデータ４１０の要素の対応する順序とを表している。

【００６９】
さまざまな実施形態では、セキュア・データ分散プログラム２００が、エンコードされたデータ４１０の分散されたストレージに関連するデータ４０１に関連付けられたメタデータを生成する。この例では、メタデータ４３１Ａは、チャンク４３０Ａに関連するメタデータを表している。同様に、メタデータ４３１ＢはチャンクＢに関連するメタデータを表しており、メタデータ４３１ＣはチャンクＣに関連するメタデータを表しており、メタデータ４３１ＤはチャンクＤに関連するメタデータを表している。一実施形態では、セキュア・データ分散プログラム２００は、元のデータ（例えば、データ４０１）に対応するデータＩＤ、チャンクＩＤ識別子、および位置識別子を示すメタデータを生成する。メタデータ４３１は、メタデータ・スキームの例を示している。図に示された例では、メタデータ４３１Ｂは「Ａ２＋２５」であり、これは、チャンク４３０ＢがデータＩＤ「Ａ」および「２」のチャンクＩＤに関連付けられており、「＋」がデータ・デリネータ（data delineator）の例であり、このデリネータの右側の情報が、エンコードされたデータの対応するチャンクを格納することに関連付けられたストレージ識別子「２５」を示すということを示している。さまざまな実施形態では、ストレージ識別子は、エンコードされたデータのチャンクを格納することが識別された、ストレージ情報１０７内のストレージＩＤまたはストレージ位置あるいはその両方に対応する（図２のステップ２０４またはステップ２０８あるいはその両方）。一実施形態では、エンコードされたデータのチャンクに関連するメタデータは、ストレージ情報１０７内に含まれ、セキュアな分散された格納のための元のデータにさらに関連付けられる。

【００７０】
本開示にはクラウド・コンピューティングに関する詳細な説明が含まれているが、本明細書において示された内容の実装は、クラウド・コンピューティング環境に限定されないと理解されるべきである。むしろ、本発明の実施形態は、現在既知であるか、または今後開発される任意のその他の種類のコンピューティング環境と組み合わせて実装できる。

【００７１】
クラウド・コンピューティングは、構成可能な計算リソース（例えば、ネットワーク、ネットワーク帯域幅、サーバ、処理、メモリ、ストレージ、アプリケーション、仮想マシン、およびサービス）の共有プールへの便利なオンデマンドのネットワーク・アクセスを可能にするためのサービス提供モデルであり、管理上の手間またはサービス・プロバイダとのやりとりを最小限に抑えて、これらのリソースを迅速にプロビジョニングおよび解放することができる。このクラウド・モデルは、少なくとも５つの特徴、少なくとも３つのサービス・モデル、および少なくとも４つのデプロイメント・モデルを含むことができる。

【００７２】
特徴は、次のとおりである。

【００７３】
オンデマンドのセルフ・サービス：クラウドの利用者は、サーバの時間およびネットワーク・ストレージなどの計算能力を一方的に、サービス・プロバイダとの人間的なやりとりを必要とせず、必要に応じて自動的にプロビジョニングすることができる。

【００７４】
幅広いネットワーク・アクセス：クラウドの能力は、ネットワークを経由して利用可能であり、標準的なメカニズムを使用してアクセスできるため、異種のシン・クライアントまたはシック・クライアント・プラットフォーム（例えば、携帯電話、ラップトップ、およびＰＤＡ）による利用を促進する。

【００７５】
リソース・プール：プロバイダの計算リソースは、プールされ、マルチテナント・モデルを使用して複数の利用者に提供される。さまざまな物理的および仮想的リソースが、要求に従って動的に割り当ておよび再割り当てされる。場所に依存しないという感覚があり、利用者は通常、提供されるリソースの正確な場所に関して管理することも知ることもないが、さらに高い抽象レベルでは、場所（例えば、国、州、またはデータセンター）を指

定できる場合がある。
【００７６】
迅速な順応性：クラウドの能力は、迅速かつ柔軟に、場合によっては自動的にプロビジョニングされ、素早くスケールアウトし、迅速に解放されて素早くスケールインすることができる。プロビジョニングに使用できる能力は、利用者には、多くの場合、任意の量をいつでも無制限に購入できるように見える。
【００７７】
測定されるサービス：クラウド・システムは、計測機能を活用することによって、サービスの種類（例えば、ストレージ、処理、帯域幅、およびアクティブなユーザのアカウント）に適した抽象レベルで、リソースの使用を自動的に制御および最適化する。リソースの使用量は監視、制御、および報告することができ、利用されるサービスのプロバイダと利用者の両方に透明性が提供される。
【００７８】
サービス・モデルは、次のとおりである。
【００７９】
ＳａａＳ（Software as a Service）：利用者に提供される能力は、クラウド・インフラストラクチャ上で稼働しているプロバイダのアプリケーションの利用である。それらのアプリケーションは、Ｗｅｂブラウザ（例えば、Ｗｅｂベースの電子メール）などのシン・クライアント・インターフェイスを介して、さまざまなクライアント・デバイスからアクセスできる。利用者は、ネットワーク、サーバ、オペレーティング・システム、ストレージ、またはさらに個々のアプリケーション機能を含む基盤になるクラウド・インフラストラクチャを、限定的なユーザ固有のアプリケーション構成設定を行う可能性を除き、管理することも制御することもない。
【００８０】
ＰａａＳ（Platform as a Service）：利用者に提供される能力は、プロバイダによってサポートされるプログラミング言語およびツールを使用して作成された、利用者が作成または取得したアプリケーションをクラウド・インフラストラクチャにデプロイすることである。利用者は、ネットワーク、サーバ、オペレーティング・システム、またはストレージを含む基盤になるクラウド・インフラストラクチャを管理することも制御することもないが、デプロイされたアプリケーション、および場合によってはアプリケーション・ホスティング環境の構成を制御することができる。
【００８１】
ＩａａＳ（Infrastructure as a Service）：利用者に提供される能力は、処理、ストレージ、ネットワーク、およびその他の基本的な計算リソースのプロビジョニングであり、ここで、利用者は、オペレーティング・システムおよびアプリケーションを含むことができる任意のソフトウェアをデプロイして実行できる。利用者は、基盤になるクラウド・インフラストラクチャを管理することも制御することもないが、オペレーティング・システム、ストレージ、デプロイされたアプリケーションを制御し、場合によっては、選択されたネットワーク・コンポーネント（例えば、ホスト・ファイアウォール）を限定的に制御する。
【００８２】
デプロイメント・モデルは、次のとおりである。
【００８３】
プライベート・クラウド：このクラウド・インフラストラクチャは、組織のためにのみ動作される。この組織またはサード・パーティによって管理することができ、オンプレミスまたはオフプレミスに存在することができる。
【００８４】
コミュニティ・クラウド：このクラウド・インフラストラクチャは、複数の組織によって共有され、関心事（例えば、任務、セキュリティ要件、ポリシー、およびコンプライアンスに関する考慮事項）を共有している特定のコミュニティをサポートする。これらの組

織またはサード・パーティによって管理することができ、オンプレミスまたはオフプレミスに存在することができる。

【００８５】

パブリック・クラウド：このクラウド・インフラストラクチャは、一般ユーザまたは大規模な業界団体が使用できるようになっており、クラウド・サービスを販売する組織によって所有される。

【００８６】

ハイブリッド・クラウド：このクラウド・インフラストラクチャは、データとアプリケーションの移植を可能にする標準化された技術または独自の技術（例えば、クラウド間の負荷バランスを調整するためのクラウド・バースト）によって固有の実体を残したまま互いに結合された２つ以上のクラウド（プライベート、コミュニティ、またはパブリック）の複合である。

【００８７】

クラウド・コンピューティング環境は、ステートレス、疎結合、モジュール性、および意味的相互運用性に重点を置いたサービス指向の環境である。クラウド・コンピューティングの中心になるのは、相互接続されたノードのネットワークを含んでいるインフラストラクチャである。

【００８８】

図５はさまざまなコンピューティング・デバイスを表しており、クラウド・コンピューティング・ノード１０（図６において説明される）は、適切なクラウド・コンピューティング・ノードの一例に過ぎず、本明細書に記載された本発明の実施形態の使用または機能の範囲に関して、いかなる制限を示唆することも意図されていない。いずれにせよ、クラウド・コンピューティング・ノード１０は、本明細書に記載された本発明の実施形態の機能のいずれかを実装すること、または実行すること、あるいはその両方を行うことができる。

【００８９】

いずれにせよ、クラウド・コンピューティング・ノード１０は、前述した機能のいずれかを実装すること、または実行すること、あるいはその両方を行うことができる。実施形態例では、クラウド・コンピューティング・ノード１０はシステム１０２を表している。さまざまな実施形態では、クラウド・コンピューティング・ノード１０は、ハードウェア要素もしくはコンピューティング・デバイスまたはその両方のハードウェアの物理的実例（例えば、ＲＩＳＣベースのサーバ６２、サーバ６３など）、または図７に関してさらに説明されるハードウェア要素、コンピューティング・デバイスの仮想化された実例（例えば、仮想サーバ７１、仮想ストレージ７２、仮想ネットワーク７３など）、あるいはその両方を表してよい。

【００９０】

図５に示されているように、一部の実施形態では、クラウド・コンピューティング・ノード１０内のコンピュータ・システム５００は、システム１０２などの汎用コンピューティング・デバイスの形態で示されている。他の実施形態では、コンピュータ・システム５００は、ストレージ１２０Ａ～ストレージ１２０Ｎなどのクラウド・コンピューティング環境５０に接続された１つまたは複数のコンピューティング・デバイスを表す。コンピュータ・システム５００は、ソフトウェアおよびデータ５１０を含んでいるシステムの例である。コンピュータ・システム５００は、プロセッサ５０１、メモリ５０２、キャッシュ５０３、永続的ストレージ５０５、通信ユニット５０７、入出力（Ｉ／Ｏ：input/output）インターフェイス５０６、および通信ファブリック５０４を含んでいる。

【００９１】

通信ファブリック５０４は、メモリ５０２、キャッシュ５０３、永続的ストレージ５０５、通信ユニット５０７、および入出力（Ｉ／Ｏ）インターフェイス５０６の間の通信を提供する。通信ファブリック５０４は、プロセッサ（マイクロプロセッサ、通信プロセッサ、およびネットワーク・プロセッサなど）、システム・メモリ、周辺機器、およびシス

テム内の任意のその他のハードウェア・コンポーネントの間で、データまたは制御情報あるいはその両方を渡すために設計された、任意のアーキテクチャを使用して実装され得る。例えば、通信ファブリック５０４は、ＰＣＩ（peripheral component interconnects）バスなどの１つまたは複数のバス、またはクロスバ・スイッチを使用して実装され得る。

【００９２】

メモリ５０２および永続的ストレージ５０５は、コンピュータ可読ストレージ媒体である。この実施形態では、メモリ５０２はランダム・アクセス・メモリ（ＲＡＭ）を含んでいる。一般に、メモリ５０２は、任意の適切な揮発性または不揮発性のコンピュータ可読ストレージ媒体を含むことができる。キャッシュ５０３は、メモリ５０２から最近アクセスされたデータ、および最近アクセスされたデータに近いデータを保持することによって、プロセッサ５０１の性能を向上させる高速なメモリである。

【００９３】

本発明の実施形態を実践するために使用されるプログラム命令およびデータは、永続的ストレージ５０５、およびキャッシュ５０３を介して各プロセッサ５０１のうちの１つまたは複数によって実行するためのメモリ５０２に格納されてよい。一実施形態では、永続的ストレージ５０５は、磁気ハード・ディスク・ドライブを含んでいる。磁気ハード・ディスク・ドライブに対する代替または追加として、永続的ストレージ５０５は、半導体ハード・ドライブ、半導体ストレージ・デバイス、読み取り専用メモリ（ＲＯＭ：read-only memory）、消去可能プログラマブル読み取り専用メモリ（ＥＰＲＯＭ：erasable programmable read-only memory）、フラッシュ・メモリ、あるいはプログラム命令またはデジタル情報を格納できる任意のその他のコンピュータ可読ストレージ媒体を含むことができる。一実施形態では、ストレージ１０３は永続的ストレージ５０５に含まれる。

【００９４】

永続的ストレージ５０５によって使用される媒体は、取り外し可能であってもよい。例えば、取り外し可能ハード・ドライブを、永続的ストレージ５０５に使用できる。その他の例としては、永続的ストレージ５０５の一部でもある別のコンピュータ可読ストレージ媒体に転送するためのドライブに挿入される、光ディスクおよび磁気ディスク、サム・ドライブ、ならびにスマート・カードが挙げられる。ソフトウェアおよびデータ５１０は、キャッシュ５０３およびメモリ５０２のうちの１つまたは複数のメモリを介して各プロセッサ５０１のうちの１つまたは複数によってアクセスするか、または実行するか、あるいはその両方を行うために、永続的ストレージ５０５に格納される。システム１０２に関して、ソフトウェアおよびデータ５１０は、データ１０４、データ・キー１０５、セキュリティ・アルゴリズム１０６、ストレージ情報１０７、セキュア・データ分散プログラム２００、セキュア・データ再構成プログラム３００、およびさまざまなプログラム（図示されていない）を含んでいる。ストレージ１２０Ａ～ストレージ１２０Ｎに関して、ソフトウェアおよびデータ５１０は、さまざまなプログラムおよびデータ（図示されていない）をそれぞれ含んでいる。

【００９５】

これらの例では、通信ユニット５０７は、システム１０２およびストレージ１２０Ａ～ストレージ１２０Ｎのリソースを含む、他のデータ処理システムまたはデバイスとの通信を提供する。これらの例において、通信ユニット５０７は、１つまたは複数のネットワーク・インターフェイス・カードを含む。通信ユニット５０７は、物理的通信リンクまたは無線通信リンクのいずれか、あるいはその両方を使用して通信を提供できる。本発明の実施形態を実践するために使用されるプログラム命令およびデータは、通信ユニット５０７を介して永続的ストレージ５０５にダウンロードされてよい。

【００９６】

Ｉ／Ｏインターフェイス５０６は、ノード１０の実例などの各コンピューティング・システムに接続されてよい他のデバイスとのデータの入力および出力を可能にする。例えば、Ｉ／Ｏインターフェイス５０６は、キーボード、キーパッド、タッチ・スクリーン、ま

たはその他の適切な入力デバイス、あるいはその組み合わせなどの、外部デバイス５０８への接続を提供してよい。外部デバイス５０８は、例えばサム・ドライブ、ポータブル光ディスクまたはポータブル磁気ディスク、およびメモリ・カードなどの、ポータブル・コンピュータ可読ストレージ媒体を含むこともできる。本発明の実施形態を実践するために使用されるソフトウェアおよびデータ５１０は、そのようなポータブル・コンピュータ可読ストレージ媒体に格納することができ、Ｉ／Ｏインターフェイス５０６を介して永続的ストレージ５０５に読み込むことができる。Ｉ／Ｏインターフェイス５０６は、ディスプレイ５０９にも接続される。

【００９７】

また、Ｉ／Ｏインターフェイス５０６は、１つまたは複数のデバイスが、コンピュータ・システム／サーバ５００または任意のデバイス（例えば、ネットワーク・カード、モデムなど）あるいはその両方と情報をやりとりできるようにし、コンピュータ・システム／サーバ５００が、１つまたは複数の他のコンピューティング・デバイスと通信できるようにする。このような通信は、Ｉ／Ｏインターフェイス５０６を介して行うことができる。さらに、コンピュータ・システム／サーバ５００は、ローカル・エリア・ネットワーク（ＬＡＮ）、一般的な広域ネットワーク（ＷＡＮ）、またはパブリック・ネットワーク（例えば、インターネット）、あるいはその組み合わせなどの１つまたは複数のネットワークと、通信ユニット５０７を介して通信することができる。図示されていないが、その他のハードウェア・コンポーネントまたはソフトウェア・コンポーネントあるいはその両方を、コンピュータ・システム／サーバ５００と併用できるということが理解されるべきである。その例として、マイクロコード、デバイス・ドライバ、冗長プロセッシング・ユニット、外部ディスク・ドライブ・アレイ、ＲＡＩＤシステム、テープ・ドライブ、およびデータ・アーカイブ・ストレージ・システムなどが挙げられるが、これらに限定されない。

【００９８】

ディスプレイ５０９は、データをユーザに表示するためのメカニズムを提供し、例えば、コンピュータのモニタであってよい。ディスプレイ５０９は、タブレット・コンピュータまたはスマートフォンのディスプレイなどの、タッチ・スクリーンとして機能することもできる。

【００９９】

図６は、クラウド・コンピューティング環境５０の例を示している。図示されているように、クラウド・コンピューティング環境５０は、クラウドの利用者によって使用されるローカル・コンピューティング・デバイス（例えば、ＰＤＡ（personal digital assistant）または携帯電話５４Ａ、デスクトップ・コンピュータ５４Ｂ、ラップトップ・コンピュータ５４Ｃ、または自動車コンピュータ・システム５４Ｎ、あるいはその組み合わせなど）が通信できる１つまたは複数のクラウド・コンピューティング・ノード１０を含んでいる。ノード１０は、互いに通信してよい。ノード１０は、１つまたは複数のネットワーク内で、本明細書において前述されたプライベート・クラウド、コミュニティ・クラウド、パブリック・クラウド、またはハイブリッド・クラウド、あるいはこれらの組み合わせなどに、物理的または仮想的にグループ化されてよい（図示されていない）。これによって、クラウド・コンピューティング環境５０は、クラウドの利用者がローカル・コンピューティング・デバイス上でリソースを維持する必要のないインフラストラクチャ、プラットフォーム、またはＳａａＳ、あるいはその組み合わせを提供できる。図６に示されたコンピューティング・デバイス５４Ａ～Ｎの種類は、例示のみが意図されており、コンピューティング・ノード１０およびクラウド・コンピューティング環境５０は、任意の種類のネットワークまたはネットワーク・アドレス可能な接続（例えば、Ｗｅｂブラウザを使用した接続）あるいはその両方を経由して任意の種類のコンピュータ制御デバイスと通信することができると理解される。加えて、クラウド・コンピューティング・ノード１０およびコンピューティング・デバイス５４Ａ～Ｎの実例は、本発明の実施形態に従って、図５に関してすでに示され、さらに詳細に説明されたコンポーネントを含んでよい。

【０１００】

図７は、クラウド・コンピューティング環境５０（図６）によって提供される機能的抽象レイヤの例示的なセットを示している。図７に示されたコンポーネント、レイヤ、および機能は、例示のみが意図されており、本発明の実施形態がこれらに限定されないということが、あらかじめ理解されるべきである。図示されているように、次のレイヤおよび対応する機能が提供される。

【０１０１】

ハードウェアおよびソフトウェア・レイヤ６０は、ハードウェア・コンポーネントおよびソフトウェア・コンポーネントを含む。ハードウェア・コンポーネントの例としては、メインフレーム６１、ＲＩＳＣ（Reduced Instruction Set Computer）アーキテクチャベースのサーバ６２、サーバ６３、ブレード・サーバ６４、ストレージ・デバイス６５、ならびにネットワークおよびネットワーク・コンポーネント６６が挙げられる。一部の実施形態では、ソフトウェア・コンポーネントは、ネットワーク・アプリケーション・サーバ・ソフトウェア６７およびデータベース・ソフトウェア６８を含む。

【０１０２】

仮想化レイヤ７０は、仮想サーバ７１、仮想ストレージ７２、仮想プライベート・ネットワークを含む仮想ネットワーク７３、仮想アプリケーションおよびオペレーティング・システム７４、ならびに仮想クライアント７５などの仮想的実体を提供できる抽象レイヤを備える。

【０１０３】

一例を挙げると、管理レイヤ８０は、以下で説明される機能を提供することができる。リソース・プロビジョニング８１は、クラウド・コンピューティング環境内でタスクを実行するために利用される計算リソースおよびその他のリソースの動的調達を行う。計測および価格設定８２は、クラウド・コンピューティング環境内でリソースが利用される際のコスト追跡、およびそれらのリソースの利用に対する請求書の作成と送付を行う。一例を挙げると、それらのリソースは、アプリケーション・ソフトウェア・ライセンスを含んでよい。ユーザ・ポータル８３は、クラウド・コンピューティング環境へのアクセスを利用者およびシステム管理者に提供する。サービス・レベル管理８４は、必要なサービス・レベルを満たすように、クラウドの計算リソースの割り当てと管理を行う。サービス水準合意（ＳＬＡ：Service Level Agreement）計画および実行８５は、今後の要求が予想されるクラウドの計算リソースの事前準備および調達を、ＳＬＡに従って行う。セキュリティ８６は、クラウドの利用者およびタスクの識別情報の検証を行うとともに、データおよびその他のリソースの保護を行う。さまざまな実施形態では、セキュリティ８６は、エンコードされてチャンク化されたデータを、異なる物理的位置で、ハードウェアおよびソフトウェア・レイヤ６０の複数のサーバまたはストレージ・システムに分散することによって、ユーザまたは企業のデータを保護する、１つまたは複数のシステムベースの機能または加入サービス（例えば、セキュア・データ分散プログラム２００およびセキュア・データ再構成プログラム３００）あるいはその両方を含む。一部の実施形態では、セキュリティ８６は、管理レイヤ８０の他の側面を利用して、異なるベンダーまたはクラウドストレージ・プロバイダのハードウェアおよびソフトウェア・レイヤ６０内のリソースに対するアクセス権限を取得し、ユーザまたは企業のエンコードされてチャンク化されたデータを格納する。

【０１０４】

ワークロード・レイヤ９０は、クラウド・コンピューティング環境で利用できる機能の例を示している。このレイヤから提供されてよいワークロードおよび機能の例としては、マッピングおよびナビゲーション９１、ソフトウェア開発およびライフサイクル管理９２、仮想クラスルーム教育の配信９３、データ解析処理９４、トランザクション処理９５、およびセキュリティ９６が挙げられる。さまざまな実施形態では、セキュリティ８６に対する代替または追加として、セキュリティ９６は、エンコードされてチャンク化されたデータを、異なる物理的位置で、ハードウェアおよびソフトウェア・レイヤ６０の複数のサーバまたはストレージ・システムに分散することによってクラウド内のデータを保護する

ためにユーザが呼び出す、セキュア・データ分散プログラム２００およびセキュア・データ再構成プログラム３００などの、１つまたは複数のソフトウェア・アプリケーションを含む。

【０１０５】

本明細書に記載されたプログラムは、アプリケーションに基づいて識別され、本発明の特定の実施形態において、そのアプリケーションに関して実装される。ただし、本明細書における任意の特定のプログラムの名前は単に便宜上使用されていると理解されるべきであり、したがって、本発明は、そのような名前によって識別されたか、または暗示されたか、あるいはその両方によって任意の特定のアプリケーションのみで使用するように制限されるべきではない。

【０１０６】

本発明は、システム、方法、またはコンピュータ・プログラム製品、あるいはその組み合わせであってよい。コンピュータ・プログラム製品は、プロセッサに本発明の態様を実行させるためのコンピュータ可読プログラム命令を有しているコンピュータ可読ストレージ媒体を含んでよい。

【０１０７】

コンピュータ可読ストレージ媒体は、命令実行デバイスによって使用するための命令を保持および格納できる有形のデバイスであることができる。コンピュータ可読ストレージ媒体は、例えば、電子ストレージ・デバイス、磁気ストレージ・デバイス、光ストレージ・デバイス、電磁ストレージ・デバイス、半導体ストレージ・デバイス、またはこれらの任意の適切な組み合わせであってよいが、これらに限定されない。コンピュータ可読ストレージ媒体のさらに具体的な例の非網羅的リストは、ポータブル・コンピュータ・ディスケット、ハード・ディスク、ランダム・アクセス・メモリ（ＲＡＭ：random access memory）、読み取り専用メモリ（ＲＯＭ：read-only memory）、消去可能プログラマブル読み取り専用メモリ（ＥＰＲＯＭ：erasable programmable read-only memoryまたはフラッシュ・メモリ）、スタティック・ランダム・アクセス・メモリ（ＳＲＡＭ：static random access memory）、ポータブル・コンパクト・ディスク読み取り専用メモリ（ＣＤ－ＲＯＭ：compact disc read-only memory）、デジタル多用途ディスク（ＤＶＤ：digital versatile disk）、メモリ・スティック、フロッピー（Ｒ）・ディスク、パンチカードまたは命令が記録されている溝の中の隆起構造などの機械的にエンコードされるデバイス、およびこれらの任意の適切な組み合わせを含む。本明細書において使用されるとき、コンピュータ可読ストレージ媒体は、それ自体が、電波またはその他の自由に伝搬する電磁波、導波管またはその他の送信媒体を伝搬する電磁波（例えば、光ファイバ・ケーブルを通過する光パルス）、あるいはワイヤを介して送信される電気信号などの一過性の信号であると解釈されるべきではない。

【０１０８】

本明細書に記載されたコンピュータ可読プログラム命令は、コンピュータ可読ストレージ媒体から各コンピューティング・デバイス／処理デバイスへ、またはネットワーク（例えば、インターネット、ローカル・エリア・ネットワーク、広域ネットワーク、または無線ネットワーク、あるいはその組み合わせ）を介して外部コンピュータまたは外部ストレージ・デバイスへダウンロードされ得る。このネットワークは、銅伝送ケーブル、光伝送ファイバ、無線送信、ルータ、ファイアウォール、スイッチ、ゲートウェイ・コンピュータ、またはエッジ・サーバ、あるいはその組み合わせを備えてよい。各コンピューティング・デバイス／処理デバイス内のネットワーク・アダプタ・カードまたはネットワーク・インターフェイスは、コンピュータ可読プログラム命令をネットワークから受信し、それらのコンピュータ可読プログラム命令を各コンピューティング・デバイス／処理デバイス内のコンピュータ可読ストレージ媒体に格納するために転送する。

【０１０９】

本発明の動作を実行するためのコンピュータ可読プログラム命令は、アセンブラ命令、命令セットアーキテクチャ（ＩＳＡ：instruction-set-architecture）命令、マシン命令

、マシン依存命令、マイクロコード、ファームウェア命令、状態設定データ、あるいは、Ｓｍａｌｌｔａｌｋ、Ｃ＋＋などのオブジェクト指向プログラミング言語、および「Ｃ」プログラミング言語または同様のプログラミング言語などの従来の手続き型プログラミング言語を含む１つまたは複数のプログラミング言語の任意の組み合わせで記述されたソース・コードまたはオブジェクト・コードであってよい。コンピュータ可読プログラム命令は、ユーザのコンピュータ上で全体的に実行すること、ユーザのコンピュータ上でスタンドアロン・ソフトウェア・パッケージとして部分的に実行すること、ユーザのコンピュータ上およびリモート・コンピュータ上でそれぞれ部分的に実行すること、あるいはリモート・コンピュータ上またはサーバ上で全体的に実行することができる。後者のシナリオでは、リモート・コンピュータは、ローカル・エリア・ネットワーク（ＬＡＮ）または広域ネットワーク（ＷＡＮ）を含む任意の種類のネットワークを介してユーザのコンピュータに接続されてよく、または接続は、（例えば、インターネット・サービス・プロバイダを使用してインターネットを介して）外部コンピュータに対して行われてよい。一部の実施形態では、本発明の態様を実行するために、例えばプログラマブル論理回路、フィールドプログラマブル・ゲート・アレイ（ＦＰＧＡ：field-programmable gate arrays）、またはプログラマブル・ロジック・アレイ（ＰＬＡ：programmable logic arrays）を含む電子回路は、コンピュータ可読プログラム命令の状態情報を利用することによって、電子回路をカスタマイズするためのコンピュータ可読プログラム命令を実行してよい。

【０１１０】

本発明の態様は、本明細書において、本発明の実施形態に従って、方法、装置（システム）、およびコンピュータ・プログラム製品のフローチャート図またはブロック図あるいはその両方を参照して説明される。フローチャート図またはブロック図あるいはその両方の各ブロック、ならびにフローチャート図またはブロック図あるいはその両方に含まれるブロックの組み合わせが、コンピュータ可読プログラム命令によって実装され得るということが理解されるであろう。

【０１１１】

これらのコンピュータ可読プログラム命令は、コンピュータまたはその他のプログラム可能なデータ処理装置のプロセッサを介して実行される命令が、フローチャートまたはブロック図あるいはその両方のブロックに指定される機能／動作を実施する手段を作り出すべく、汎用コンピュータ、専用コンピュータ、または他のプログラム可能なデータ処理装置のプロセッサに提供されてマシンを作り出すものであってよい。これらのコンピュータ可読プログラム命令は、命令が格納されたコンピュータ可読ストレージ媒体がフローチャートまたはブロック図あるいはその両方のブロックに指定される機能／動作の態様を実施する命令を含んでいる製品を備えるように、コンピュータ可読ストレージ媒体に格納され、コンピュータ、プログラム可能なデータ処理装置、または他のデバイス、あるいはその組み合わせに特定の方式で機能するように指示できるものであってもよい。

【０１１２】

コンピュータ可読プログラム命令は、コンピュータ上、その他のプログラム可能な装置上、またはその他のデバイス上で実行される命令が、フローチャートまたはブロック図あるいはその両方のブロックに指定される機能／動作を実施するように、コンピュータ、その他のプログラム可能なデータ処理装置、またはその他のデバイスに読み込まれてもよく、それによって、一連の動作可能なステップを、コンピュータ上、その他のプログラム可能な装置上、またはコンピュータ実装プロセスを生成するその他のデバイス上で実行させる。

【０１１３】

図内のフローチャートおよびブロック図は、本発明のさまざまな実施形態に従って、システム、方法、およびコンピュータ・プログラム製品の可能な実装のアーキテクチャ、機能、および動作を示す。これに関連して、フローチャートまたはブロック図内の各ブロックは、指定された論理機能を実装するための１つまたは複数の実行可能な命令を備える、命令のモジュール、セグメント、または部分を表してよい。一部の代替の実装では、ブロ

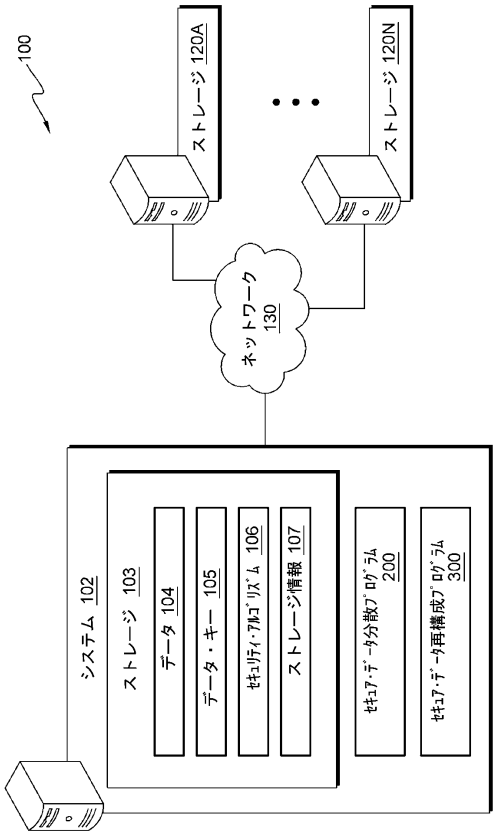
ックに示された機能は、図に示された順序とは異なる順序で発生してよい。例えば、連続して示された２つのブロックは、実際には、含まれている機能に応じて、実質的に同時に実行されるか、または場合によっては逆の順序で実行されてよい。ブロック図またはフローチャート図あるいはその両方の各ブロック、ならびにブロック図またはフローチャート図あるいはその両方に含まれるブロックの組み合わせは、規定された機能または動作を実行するか、または専用ハードウェアとコンピュータ命令の組み合わせを実行する専用ハードウェアベースのシステムによって実装され得るということにも注意する。

【０１１４】

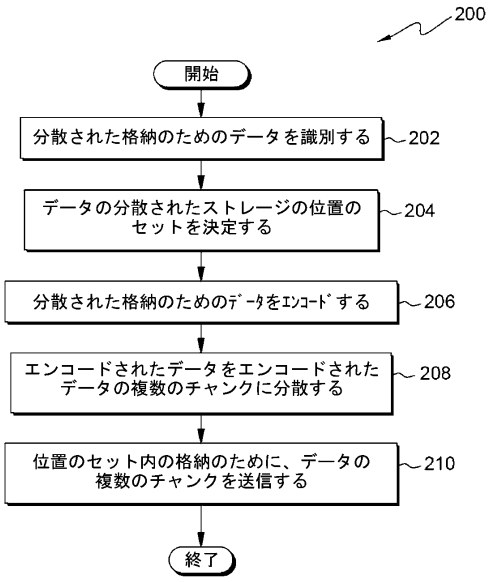
本発明のさまざまな実施形態の説明は、例示の目的で提示されているが、網羅的であることは意図されておらず、開示された実施形態に制限されない。本発明の範囲および思想を逸脱することなく多くの変更および変形が可能であることは、当業者にとって明らかである。本明細書で使用された用語は、実施形態の原理、実際のアプリケーション、または市場で見られる技術を超える技術的改良を最も適切に説明するため、または他の当業者が本明細書で開示された実施形態を理解できるようにするため選択されている。

【図面】

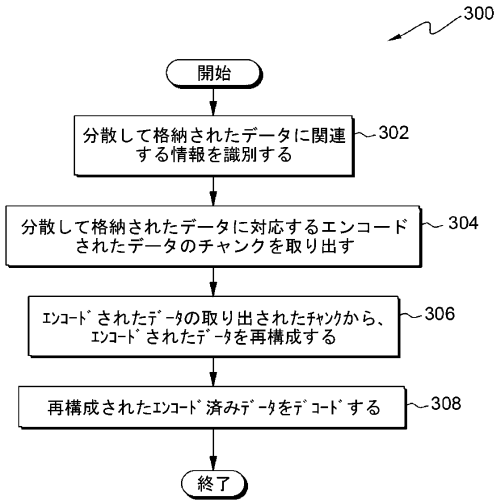
【図１】


100
ストレージ 120A
ストレージ 120N
ネットワーク 130
システム 102
ストレージ 103
データ 104
データ・キー 105
ｾｷｭﾘﾃｨ･ｱﾙｺﾞﾘｽﾞﾑ 106
ストレージ情報 107
ｾｷｭｱ･ﾃﾞｰﾀ分散ﾌﾟﾛｸﾞﾗﾑ 200
ｾｷｭｱ･ﾃﾞｰﾀ再構成ﾌﾟﾛｸﾞﾗﾑ 300

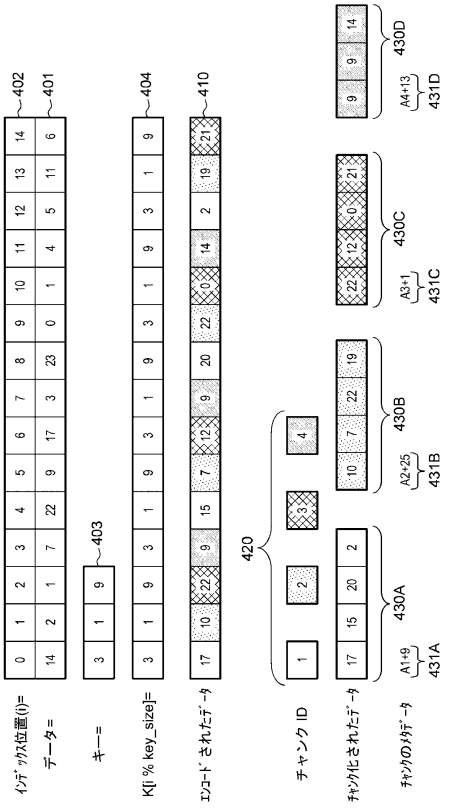
【図２】


200
開始
分散された格納のためのデータを識別する
202
データの分散されたストレージの位置のセットを決定する
204
分散された格納のためのﾃﾞｰﾀをｴﾝｺｰﾄﾞする
206
エンコードされたデータをエンコードされたデータの複数のチャンクに分散する
208
位置のセット内の格納のために、データの複数のチャンクを送信する
210
終了

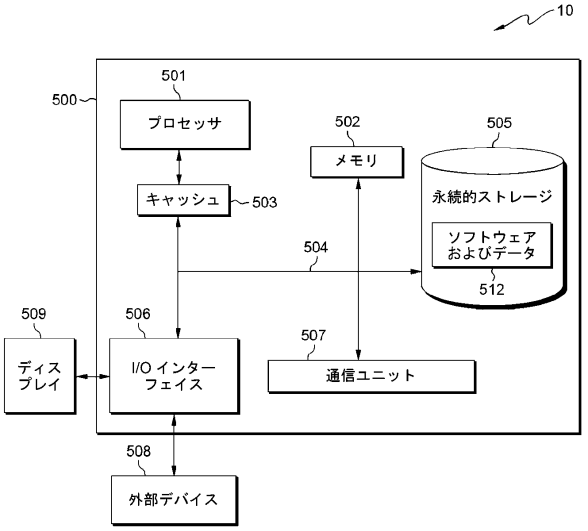
【図３】


300
開始
分散して格納されたデータに関連する情報を識別する
302
分散して格納されたデータに対応するエンコードされたデータのチャンクを取り出す
304
ｴﾝｺｰﾄﾞされたﾃﾞｰﾀの取り出されたﾁｬﾝｸから、ｴﾝｺｰﾄﾞされたﾃﾞｰﾀを再構成する
306
再構成されたｴﾝｺｰﾄﾞ済みﾃﾞｰﾀをﾃﾞｺｰﾄﾞする
308
終了

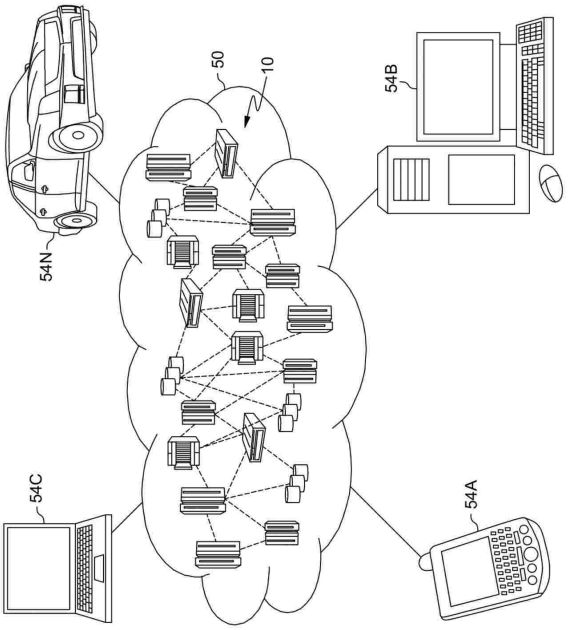
【図４】


402
401
403
404
410
420
430A
430B
430C
430D
431A
431B
431C
431D
A1+9
A2+25
A3+1
A4+13
ｲﾝﾃﾞｯｸｽ位置(i)=
ﾃﾞｰﾀ=
ｷｰ=
K[i % key_size]=
ｴﾝｺｰﾄﾞされたﾃﾞｰﾀ
ﾁｬﾝｸ ID
ﾁｬﾝｸ化されたﾃﾞｰﾀ
ﾁｬﾝｸのﾒﾀﾃﾞｰﾀ

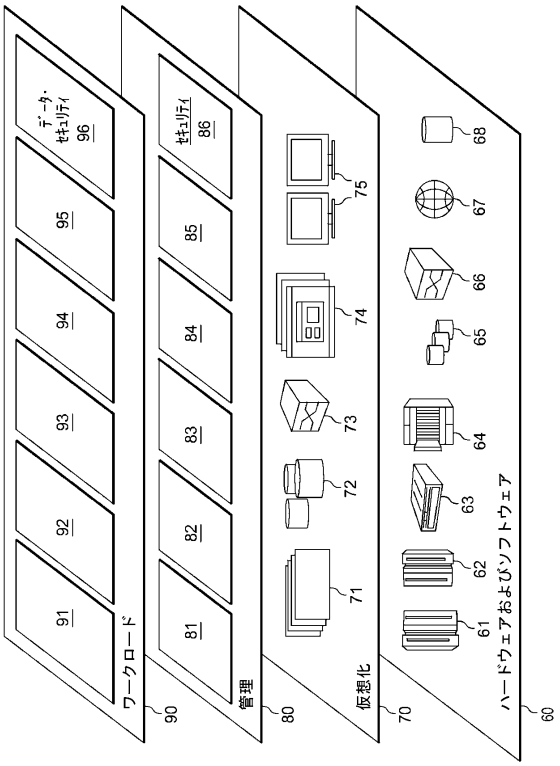
【図５】


10
500
501
プロセッサ
502
メモリ
505
永続的ストレージ
キャッシュ
503
ソフトウェア
およびデータ
504
512
509
506
507
ディス
プレイ
I/O インター
フェイス
通信ユニット
508
外部デバイス

【図６】


50
10
54B
54N
54C
54A

【図７】


ワークロード
90
管理
80
仮想化
70
ハードウェアおよびソフトウェア
60
91
92
93
94
95
96
81
82
83
84
85
86
61
62
63
64
65
66
67
68
71
72
73
74
75

フロントページの続き

(72)発明者　モル、ジョージス－ヘンリー
フランス ビオ０６４１０ アルテ・パーク・キャンバス アベニュー・ルマニール・ゼットエーシー・ストリート バット・シー９６５
(72)発明者　オウドット、オリビエ
フランス ビオ０６４１０ アルテ・パーク・キャンバス アベニュー・ルマニール・ゼットエーシー・ストリート バット・シー９６５
(72)発明者　カプラン、フィリッペ
フランス ビオ０６４１０ アルテ・パーク・キャンバス アベニュー・ルマニール・ゼットエーシー・ストリート バット・シー９６５
(72)発明者　ノデット、ザビアー
フランス ビオ０６４１０ アルテ・パーク・キャンバス アベニュー・ルマニール・ゼットエーシー・ストリート バット・シー９６５

審査官　金沢 史明
(56)参考文献　特開２００４－１７１２０７（ＪＰ，Ａ）
米国特許出願公開第２０１８／００６０６１２（ＵＳ，Ａ１）
国際公開第２００８／０５６６６７（WO，Ａ１）
特開２００９－１２２７３１（ＪＰ，Ａ）
国際公開第２０１５／０８９６２３（WO，Ａ１）
特開２００４－２６５１９４（ＪＰ，Ａ）
特開２０００－１２２８６１（ＪＰ，Ａ）
特開２０１１－１１９９８５（ＪＰ，Ａ）
(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｇ０９Ｃ　１／００
Ｇ０６Ｆ　２１／６２
Ｇ０６Ｆ　１６／１８２