

① 集中データベースと分散データベース

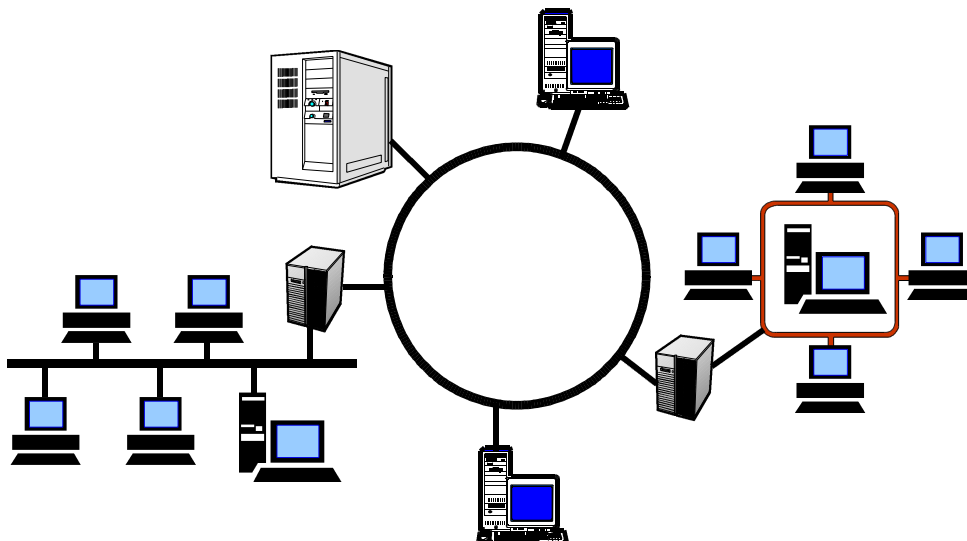
② 集中型データベース

データを集中管理し、システムの開発や保守、運用管理などのコストを削減することができるため、従来のデータベースの主流であった。

③ 集中型データベースの問題点

- ㊦ データベースの障害がシステム全体に及ぶ。
- ㊧ 特定部門の要求に対して敏速に対応できない。
- ㊨ 発生するデータは通信回線を通して集中処理するので、データ通信のための費用がかかる。
- ㊩ データベースが巨大化し、費用や要員が増大する。

④ 分散データベース



集中データベースの問題を解消するために、データベースを分散し、複数箇所のデータベースをあたかも一つのデータベースとして利用できる分散データベースが生まれた。分散データベースはデータベースとネットワーク環境の進展の結果、生まれたものである。

工場や営業所、研究所など、地理的に離れた場所の複数のコンピュータシステムに接続されたデータベースを、論理的に一つのデータベースとしてアクセスできるようにしたデータベースである。異機種間で共通のSQLによって、他の分散システムにあるデータベースも自分のところのデータベースと同じように利用できる。

④ 分散データベースの利点

- ㊦ 各部門の利用者自身が必要な情報の検索や編集を行うことができ、操作も簡単である
- ㊧ 企業を取り巻く業態の変化に対応しやすい。
- ㊨ 各部門ごとに独立して処理が行われるので、部門内の要求をそのままシステムに組み込むことができる。
- ㊩ 操作場所にデータベースがあるので、応答時間が短い。
- ㊪ 各地区ごとに頻繁に利用するデータベースをもつことによって、通信費用の削減ができる。
- ㊫ データベースに障害が発生しても、他のデータベースを使用することができるので、危険性を分散することができる。
- ㊬ 他のデータベースにアクセスするときにも、利用者はデータベースがどこにあるのかということを全く意識しなくてもよい。
- ㊭ 各サイトが独立しているので拡張性に富む。

⑤ 分散データベースの問題点

- ㊦ セキュリティ管理やパスワード管理などの運用管理が難しい。
- ㊧ データベースの分散により、プログラムが分散してしまう。
- ㊨ 部門固有の機能などが追加されるため、プログラム全体のバージョン管理が難しくなる。
- ㊩ 部門単位、個人単位にプログラムが開発されるため、同じようなプログラムを重複して作ってしまう恐れがある。
- ㊪ 全社的な処理を行う場合に、データ通信のための時間・費用がかかる。
- ㊫ 一括処理がしにくい。

② 分散データベースの透過性

㊦ 透過性

分散データベースのもつ複雑な機能をユーザが意識することなく取り扱うことができる基本的な条件が透過性である。資源位置に対する透過性、データ移動に対する透過性、データ分割に対する透過性、データ重複に対する透過性、障害に対する透過性などがある。

㊧ 分散データベースの透過性の種類

㊦ 資源位置に対する透過性

ユーザがデータベースの位置を意識せずに利用できることである。データベース管理システムは、データとデータベースサイトとの対応をデータディクショナリ／ディレクトリに保持している。

① データ移動に対する透過性

運用上の都合や性能上の問題でデータの格納サイトが変更されても、ユーザ側のアプリケーションや操作方法に影響がないことである。データディクショナリ／ディレクトリの更新を前提にする。資源位置の透過性とデータ移動に対する透過性は一体のものである。

② データ分割に対する透過性

一つのデータが複数のサイトに分割されて格納されていても、ユーザがそれを意識することなく利用できることである。水平分割は、一つのデータの集まりを水平方向に分割する。レコード単位に分割し、複数サイトに格納する。各サイトに格納されたそれぞれを水平フラグメントという。垂直分割は、一つのデータの集まりを垂直方向に分割する。項目単位に分割し、複数のサイトに格納する。各サイトに格納されたそれぞれを垂直フラグメントという。

③ データ重複に対する透過性

一つのデータが複数のサイトに重複されて格納されていても、ユーザがそれを意識することなく利用できることである。論理的には一つのデータを物理的に重複させて複数のサイトにもつことである。実現する機能は、データディクショナリ／ディレクトリである。更新処理の場合、あるデータが更新されると、重複データも自動的に即時あるいは一定時間経過後、もれなく更新する仕組みである。データ形式もデータ内容も全く同じものをわざわざ重複させる意図は、検索処理における通信費の削減や信頼性の向上にある。

④ 障害に対する透過性

あるサイトで起こった障害をユーザが意識することなく、データベースが利用できることである。障害が起こったサイトでのサービスは停止するが、それ以外のサイトにおけるサービスは続行する場合や障害が起こったサイトの代替サービスを別のサイトが行うことにより、システム全体のサービス低下を最低限度に抑える場合がある。後者の方が障害に対する透過性が高いという。複数サイトでのデータの重複が不可欠である。

⑤ データモデルに対する透過性

各サイトにおけるデータベース管理システムが異なったデータモデルであっても、ユーザは意識することなく利用できることである。このようなデータベースをマルチデータベースという。

③ RDA (遠隔データベースアクセス)

① RDAとは

異機種間でも共通のSQLによって、他の分散システムにあるデータベースも、自分のところのデータベースと同様に利用できる仕組みのことである。

⑥ RDAの仕組み

㊦ ユーザは、SQL文でデータ処理要求を出す。

分散DBMSは、あらかじめ登録されているディレクトリから、参照すべきデータベースの所在場所を探す。

① 分散DBMSは、参照すべきデータベースを持っているシステムに対して、ユーザが指定したSQL文に従って、データ処理を要求する。

RDAクライアントは、データ処理を要求する機能と結果を受け取る機能を持っている。
RDAサーバは、データ処理を受理する機能と結果を送信する機能を持っている

㊧ RDAサーバは、RDAクライアントの要求に基づいて、データベース言語を使ってデータ処理を行う。

② RDAサーバは、RDAクライアントに処理結果を返す。

㊨ 分散DBMSは、ユーザに処理結果を返す。

④ 2相コミットメント

① 2相コミットメントとは

データベースの更新を第一相と第二相の2回のコミュニケーションで行う方法で、データの整合性を確保する。このような分散データベースの整合性の確保の方式を2相コミットという。

異常時には、すべてのデータベースを更新するか更新前の状態に戻すことによって、データベースの整合性を確保する。

② 正常な処理

㊦ PREPARE

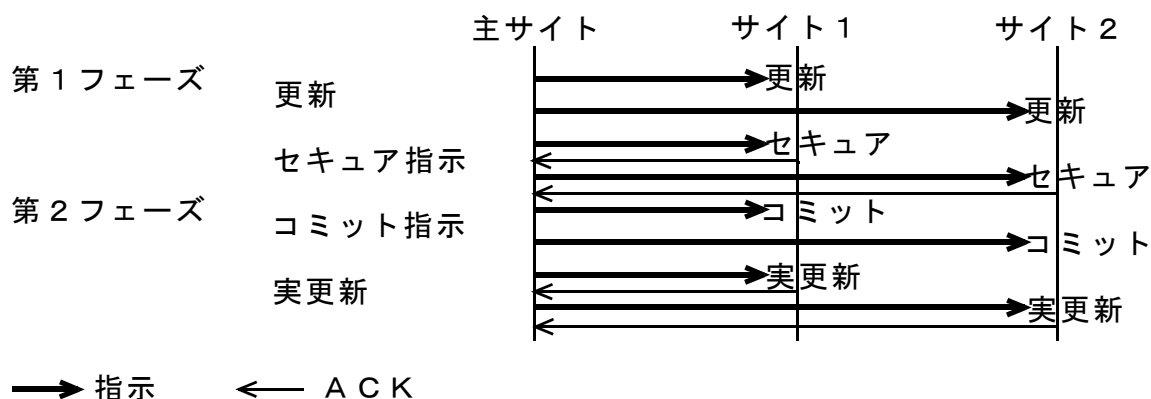
第一相で同期要求元の主サイトがサイト1、サイト2に対して、データベース更新の保証処理を依頼する。これをPREPARE という。

① COMMIT

第二相で要求元の主サイトがすべてのデータベースの“OK”を受け取って、サイト1、サイト2に実際の更新を依頼する。これをCOMMITという。

③ 異常処理

- ⑦ 異常処理は、第一相で同期の要求元の主サイトが、データベース更新の保証処理を依頼した結果、どれか一方所のサイトから“NO”の返事を受け取る場合である。
- ⑧ 同期の要求元の主サイトは、サイト1、サイト2に対して、データベースを更新前の状態に戻す依頼をする。これをROLLBACKという。



⑤ クライアント／サーバ間の通信

① C/S型システムの通信コスト削減上の問題点

- ⑦ ネットワーク上のデータ転送に伴う通信時間の問題
ネットワーク媒体の速度に依存する。
- ⑧ データ通信に伴うプロトコルハンドリングに必要な通信処理時間
データ通信回線が低速で信頼性が低い時代は転送データの削減が課題であったが、ネットワークが高速になるにつれ、通信回数の削減が重要になってきた。

② 標準のSQLでの通信回数

アプリケーションとDBMSの間では1レコード単位でしかやり取りができないため、C/S型では通信回数が増加する。1レコードの追加にC/S間では最低2回の通信が交わされる。1000レコードの追加には最低2000回の通信が交わされる。DBMSの中には複数のレコードを一括して取り扱うように拡張したものもあり、1回当たりのデータ量は増加するが、通信回数は減少し、全体の処理時間は短縮する。

③ ストアードプロシジャ

ユーザの定義したデータベースアクセス処理を1つの処理単位としてデータベース内に登録して、アプリケーション側からこれを起動できるようにしたものである。通常、トランザクションは複数のSQLのシーケンスである。

複数のレコードの送受信だけでなく、複数のSQL文を1回の送信でサーバ上のDBMSに送付し処理させると、通信回数を大幅に削減できる。複数のSQL文を処理できるように、条件判断、ループ、変数などの機能を付加したプログラム言語が利用される。

ストアードプロシジャはこのプログラム言語とSQLで作成されたプログラムが一体化されたプログラムの一種である。ストアードプロシジャはそれ自体1つのトランザクションまたはひとまとまりのデータを扱う手続きとして定義する。

④ ストアードプロシジャの効果

㉞ C/S間の通信回数の削減

複数のSQL文からなる手続きを1回で呼び出して実行するために、C/S間の通信回数を大幅に削減できる。

㉟ システム全体に共通なロジックの活用

システム全体に共通なロジックをストアードプロシジャとして格納でき、アプリケーション毎の勝手なデータアクセスを許さないシステムを実現できる。

㊱ メモリの節約

複数のアプリケーションが共通のストアードプロシジャを共有するために、メモリを節約できる。

㊲ 生産性とセキュリティの向上

オブジェクト指向的な考え方に基づいて、あるデータに対する処理をストアードプロシジャ呼び出しに限定できるために、生産性とセキュリティの向上につながる。

㊳ アプリケーションの品質向上

アプリケーション全体の整合性を保ちやすく、品質向上につながる。

⑥ オブジェクト指向データベース

㉠ オブジェクト指向データベース

㉞ オブジェクト指向データベースとは

データとその処理手続きを一体化(カプセル化)したものをオブジェクトと呼ぶ。階層データベースやネットワークデータベース、関係データベースが格納する対象は、数値や文字を中心としたデータであるが、オブジェクト指向データベースの格納対象は、オブジェクトである。このオブジェクトをディスク上に記録し、管理する。

① OODBMS

オブジェクトの生成や検索、追加、削除、更新を行うことができる仕組みをOODBMSという。OODBMSによって収集されたデータの集合がオブジェクト指向データベースである。

⑥ オブジェクト指向データベース管理システムの主要機能

- ㊦ データベース中の永続的なオブジェクトをプログラムから直接読み込む機能
- ① プログラムからデータベースにオブジェクトを直接書き込む機能
- ㊵ 複数のプログラムやユーザがデータベースの中のオブジェクトを共有し、同時にアクセスできる機能
- ㊥ 索引などの高速なアクセス機能
- ㊦ ある条件を満たすオブジェクトを一括的に検索するための問い合わせ機能

⑦ オブジェクト指向データベースの特徴

- ㊦ 格納できるデータの種別は広範囲である。
数値や文字の他に、静止画像や動画、図形、音声などがある。
- ① 扱えるデータ構造の種類が多い。
一つのデータベースの中で混在できる。複雑な構造のデータも扱える。
- ㊵ データと手続きをカプセル化して一緒に格納できる。
- ㊥ データベースの操作をユーザが意識する必要がない。
使ったオブジェクトをOODBMSが終了時にデータベースに格納する。自動的に格納されるオブジェクトを永続的オブジェクトといい、格納しないオブジェクトを一時的オブジェクトという。永続的か一時的かはユーザが決める。
- ㊦ 一つのオブジェクトは複数の参照オブジェクトIDを持つことができる。
 $n : m$ の関係をデータベース上に容易に実現できる。OODBMSがオブジェクトIDをつけるのでユーザは意識する必要がない。
- ㊦ クラスとクラスの間継承という関連をつけておくと、属性が引き継がれる。
- ㊥ 応用プログラムはオブジェクトを組み合わせ使用することができる。
- ㊦ 利用者はあらかじめ決められた形式のメッセージを送ることによって処理要求を行う。

⑧ オブジェクト指向データベース管理システムの特徴

- ㊦ オブジェクト識別子
生成されたオブジェクトには、そのオブジェクト固有の識別子が割り当てられる。
- ① 複合オブジェクト
オブジェクトの属性値として、ほかのオブジェクトの識別子を値として持たせることができる。

㊤ カプセル化

データベースの中のデータとそのデータに適用できる手続きを一体化して管理できる。

㊦ クラス階層と継承

オブジェクトをその種類に応じて分類するためにクラス階層が利用できる。継承や検索に利用できる。

⑦ マルチメディアデータベース

㊱ マルチメディアデータベース

マルチメディア データベースは、大量のファイルの整理、管理に最適なデータベースであり、ファイル形式を問わず、画像や、書類などのあらゆる形式のファイルを一括して管理できる。文字や数値に加えて画像や音声などのマルチメディアデータを扱うことができ、利用者に対してメディアのデータ構造を意識させることなく、統一されたユーザインタフェースを提供することができる。必要な資料や情報を、必要な時に簡単に、素早く探し出せるようになり、業務は大きく効率化される。

㊲ マルチメディアデータベースの機能

㊴ 複雑で巨大なデータ構造が扱える。

データベース管理システム自身がデータ構造を定義し、その構造に沿って検索や部分変更が行える。

㊵ データの時間的な操作・検索に対応できる。

動画や音声の、早回し・スロー再生・一時停止などの機能を実現する。

⑧ データウェアハウス

㊱ ビッグデータ

㊴ ビッグデータとは

ビッグデータは、インターネットの普及とIT技術の進化によって生まれたもので、これまで企業が扱ってきた以上に、大容量、多様なデータを扱う仕組みであり、データ量、データの種類、変化する頻度という3つの特性を持っている。

㊵ 扱うデータの種類

数値や文字列といったデータだけでなく、文章、音声、動画といったマルチメディア・データが含まれる。さらに、電子メールのデータやXMLデータ、各種センサーや機器から発

せられるデータや通信ログのように頻度が非常に多いデータも含まれる。しかも、社内だけでなく、インターネット上の社外にある場合も多い。

㉟ 従来の大容量データとビッグデータの特性の差異

従来の大容量データは、構造化データ中心であり、データ更新の頻度も月単位など変化頻度はそれほど高くなかった。それに対して、ビッグデータでは、大容量＋非構造化データ＋高頻度という新たな組み合わせの特性を有している。

⑥ データウェアハウス

㊲ データウェアハウス

データウェアハウスは、企業戦略の立案や意思決定などに利用するため、組織内の各種データを1カ所に集約する仕組みであり、データの倉庫である。集められるデータは、売上伝票などの生データであり、製造、物流、販売など部門ごとに蓄積していた各種データを集めて、それを使用して相関分析などを行うために用いる。

㊳ データの格納要領

データウェアハウスは、雑多なデータを単にそのまま記録するだけでなく、様々な情報システムのデータを集めて同じ意味のデータは横断的に扱えるようにするなど統合し、データの主題ごとに整理し直して格納する。また、データは発生した時系列に記録していき、更新や削除は行わず、過去のある時点の個別のデータを参照できるように格納される。データウェアハウスに集積された過去のデータを元に、抽出や分析、比較などを行なって、計画立案や意思決定などのために有用な知見を得るために構築されることが多い。

㉞ データベースとの差異

データベースは、現在進行中の状態のデータを格納し、即座に取り出せるように保管する。容量の関係や処理速度の問題で、一定期間が過ぎたデータは削除され、ビジネス業務を円滑に進めることが前提のシステムであるから、日次、月次、年次などとデータ算出のサイクルがあらかじめ定めて管理する。データウェアハウスは、データの倉庫として情報を時系列に保管する。現在は使わないものを集約し、格納するのがデータウェアハウスの役割である。

⑦ データマイニング

㊲ データマイニングとは

データマイニングは、データウェアハウスなどに蓄積された膨大な量のデータを分析し、データの中に隠れている有用な規則性を見出すことである。経営やマーケティングの分野で利用されている。

㊳ データマイニングでできること

経営戦略や技術開発戦略、新製品のマーケティング戦略を検討する際、根拠となるデータを得る等の目的で、データマイニングを実施することが多い。戦略を検討する場合に次のことを実行する。

- ① データの分類を行う
- ② データの関連性を見つけ出す
- ③ 事象の発生確率を予測する

㉔ データマイニングの手法

① マーケット・バスケット分析

マーケット・バスケット分析は、データ同士の関係性を分析するもので、どの商品とどの商品をどのような顧客が同時に購入したかを分析する手法である。コンビニのPOSデータの分析でよく使われる。

② クラスタ分析

クラスタ分析は、データを分類する分析手法で、階層クラスタ分析と非階層クラスタ分析の2種類がある。階層クラスタ分析は、最も似ている組み合わせから順にまとめていく方法である。非階層クラスタ分析は階層構造を持たないため、たくさんのデータがあっても分析が可能なことから、ビッグデータの分析に適している。

③ ロジスティック回帰分析

ロジスティック回帰分析は、発生確率を予測する手法で、マーケティング施策に対する顧客の反応を改善するためなどに使われる。

㉕ データマート

データマートはデータウェアハウスから利用者の目的に合わせて抽出し、利用者が扱いやすいようにデータの要約化などの加工を施した小規模なデータベースである。データウェアハウスが企業全体の情報を対象としたものであるのに対して、部門レベルの利用目的に特化したものである。

㉖ ETLツール

㉖-1 ETLツールとは

ETLツールは、基幹系システムから業務データを抽出、加工し、データウェアハウスに送る作業を行うツールである。データウェアハウスのデータ品質向上のため、ETLツールを導入する企業が増加している。

㉖-2 ETLツールの機能

- ① 企業の基幹系で使用されているデータベースにアクセスし、データを抽出する。

- ② FTPやHTTP、SOAPなどのプロトコル、APIでの接続に対応しており、クラウドサービスとの接続や、ビッグデータの取り込みに対応できる。
- ③ 種々のファイル形式に対応することで効率的な分析が行なえる。
- ④ 多くの企業で導入されているデータ連携ソフトウェアとの接続機能も標準で用意されている。
- ⑤ インターフェースやロジックを標準化し開発量を削減する「部品化」機能をもっている。

例題演習

分散型データベースシステムモデルで、複数のデータベース間の更新の同期をとるために用いられる方法はどれか。

- ア 2相コミット
- イ 時刻印法
- ウ 正規化
- エ ロールバック

解答解説

分散データベースの2相コミットメントに関する問題である。

アの2相コミットメントは、ネットワークに接続した複数のコンピュータに分散して配置したデータベースを更新するときに、更新処理の同期をとるための仕組みである。求める答えはアとなる。

イの時刻印法は、同時実行制御の一手段で、トランザクションの開始時刻やデータの読取り・書出しの時刻を記録し、その比較によってデータの更新順序を確保する方法である。

ウの正規化は、データの冗長性を排除して、データの一貫性と整合性を図ることである。

エのロールバックは、ある処理で障害が発生したとき、その処理の更新内容を元に戻すことで、データベースやファイルを復旧する手段である。

例題演習

クライアントサーバ型システムにおいて、サーバ側で処理するのが最も適切な機能はどれか。

- ア 出力データの表示処理
- イ データベースの更新処理
- ウ 入力されたデータの形式チェック
- エ ブルダウンメニューの表示処理

解答解説

クライアントサーバシステムのサーバの役割に関する問題である。

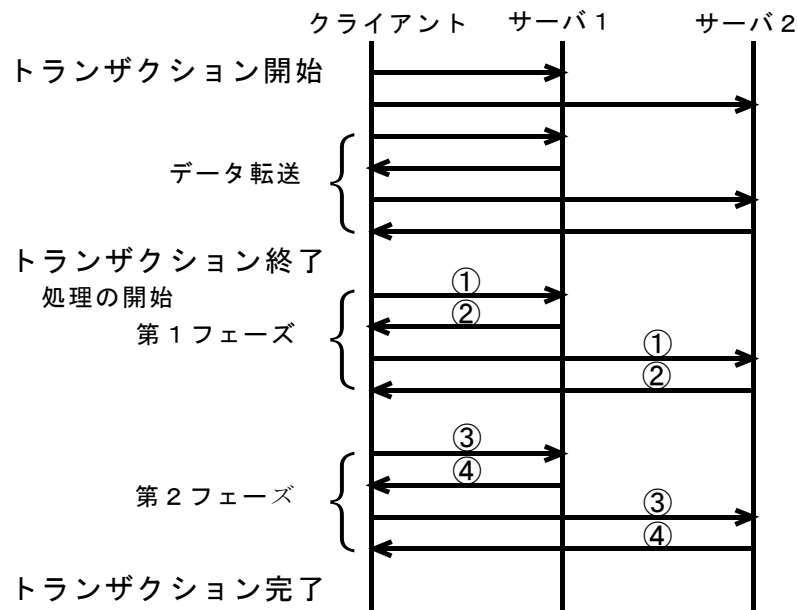
クライアントサーバシステムは、情報や資源を一元的に管理し、クライアントとサーバが連携することで、一つのアプリケーションプログラムが効率的に実行する方式である。情報や資源を提供する役割のハードやソフトをサーバ、サーバに要求を出して情報や資源を利用するハードやソフトをクライアントと呼ぶ。

アの出力データの表示処理、ウの入力されたデータの形式処理、エのプルダウンメニューの処理は、各クライアントで独自に処理するものであり、サーバが行うと大幅なネットワークの負荷増になる。

イのデータベースの更新処理は、複数のクライアントが関係するためサーバで実行する方が好ましい。求める答えはイとなる。

例題演習

図は、2相コミットメントプロトコルにおける正常処理の流れを示している。①～④の組合として適切なものはどれか。



	①	②	③	④
ア	コミット可否 問合せ	コミット可応答	コミット実行 指示	コミット実行 応答
イ	コミット実行 指示	コミット実行 応答	データベース 更新指示	データベース 更新応答
ウ	ジャーナル取得 指示	ジャーナル取得 応答	コミット実行 指示	コミット実行 応答
エ	データベース 更新指示	データベース 更新応答	メッセージ送信 指示	メッセージ送信 応答

解答解説

2相コミットメントに関する問題である。

- ① 第1フェーズは両サイトにコミットの可否を問い合わせる
- ② 両サイトへのコミット可応答
- ③ 両サイトがOKならば、両サイトにコミット実行を指示する

④ コミット実行応答でトランザクション処理が完了する。

求める答えはアとなる。

例題演習

データベースサーバを利用したクライアントサーバシステムにおいて、大量のSQL文が発生することによってクライアントとサーバ間の通信負荷が問題となった。このときの解決策として、適切なものはどれか。

ア インデックスの見直し

イ ストアドプロシージャ機能の利用

ウ データベースの再編成

エ 動的SQLの利用

解答解説

ストアドプロシージャに関する問題である。

アのインデックスはデータベースの検索や並べ替えを高速に行うために使用する索引である。キーワードがどのレコードにあるかを示したインデックスを作成することによって行う。

イのストアドプロシージャはクライアント／サーバー型のデータベース・システムで、処理を高速化するための手法である。利用頻度の高いプロシージャをあらかじめサーバー上に用意し、サーバーにあるデータベースにアクセスする際に、クライアントはプロシージャを呼び出すだけで済むようにしておく。これによりクライアント／サーバー間の通信回数や通信データ量を減らせるため、データベース・システム全体の処理スピードが向上する。求める答えはイとなる。

ウのデータベースの再編成はデータの追加や削除、変更によって記憶効率が低下し、アクセス速度が低下した場合に改善のために行う手段である。

エの動的SQLはホスト言語の実行時にSQL文を生成する方法で、パフォーマンスでは静的SQLに比して劣るが、柔軟性の高いアプリケーションを開発できる利点がある。プログラムで実行するSQL文や検索条件がプログラム作成時に決まらない場合に使用できる。

例題演習

分散データベースの透過性の説明として、適切なものはどれか。

ア クライアントのアプリケーションプログラムは、複数のサーバ上のデータベースをアクセスする。アプリケーションプログラムは、データベースがあたかも一つのサーバ上で稼働しているかのようにアクセスできる。

イ クライアントのアプリケーションプログラムは、複数のサーバ上のデータベースをアクセスする。アプリケーションプログラムはどのサーバ上のデータベースをアクセスするのかを知っている必要がある。

ウ 複数のクライアントのアプリケーションプログラムが、一つのサーバ上のデータベースを共有してアクセスする。

エ 複数のクライアントのアプリケーションプログラムは、一つのサーバ上のデータベースを、サーバ上のアプリケーションプログラムを介してアクセスする。

解答解説

分散データベースの透過性に関する問題である。

分散データベースの透過性は、分散データベースのもつ複雑な機能をユーザが意識することなく取り扱うことができる基本的な条件である。資源位置に対する透過性、データ移動に対する透過性、データ分割に対する透過性、データ重複に対する透過性、障害に対する透過性などがある。位置の透過性は資源の物理的な位置を意識しなくても利用できることであり、アクセス透過性はローカルファイルとリモートファイルを意識することなく、同一操作方法でアクセスできることである。アの説明が資源位置の透過性である。求める答えはアとなる。

イの透過性は位置を知る必要がない。ウは透過性はデータベースの共有ではない。エのサーバ上のアプリケーションを介してアクセスしない。

例題演習

ビッグデータの活用例として、大量のデータから統計学的手法などを用いて新たな知識（傾向やパターン）を見つけ出すプロセスはどれか。

ア データウェアハウス

イ データディクショナリ

ウ データマイニング

エ メタデータ

解答解説

データマイニングに関する問題である。

アのデータウェアハウスは、意思決定を支援するために、全社規模で発生する大量データを蓄積したデータベースである。

イのデータディクショナリは、データの性質などの定義を保存する辞書である。管理するデータ、利用者、プログラムに関する情報、それらの関係を保持するデータの集合体である。

ウのデータマイニングは、データウェアハウスなどに蓄積された膨大な量のデータを分析し、統計学的手法などを用いて有用な規則性や新たな知識を見いだすプロセスである。求める答えはウとなる。

エのメタデータは、検索の対象となるデータを要約した情報のことである。

例題演習

データウェアハウスに業務データを取り込むとき、データを抽出して加工し、データベースに書き出すツールはどれか。

ア ETLツール

イ OLAPツール

ウ データマイニングツール

エ 統計ツール

解答解説

データウェアハウスに関する問題である。

アのETLツールは、企業の基幹系システムなどに蓄積されたデータを抽出し、データウェアハウスなどで利用しやすい形に加工し、対象となるデータベースに書き出す機能を有したツ

ールである。また、これら一連の処理を支援するソフトウェアを指す。求める答えはアとなる。

イのOLAPツールは、会計管理や販売管理、顧客管理、在庫管理、購買管理、生産管理といった、企業の業務アプリケーションからデータを取り出し、傾向や特徴を複数の視点で分析するソフトウェアである。

ウのデータマイニングは、小売店の販売データや電話の通話履歴、クレジットカードの利用履歴など、企業に大量に蓄積されるデータを解析し、その中に潜む項目間の相関関係やパターンなどを探し出す技術である。

エの統計ツールは、統計解析のための各種ソフトウェアである。

例題演習

データマイニングの説明はどれか。

- ア 検索効率を高めるために、大規模データベースの一部をあらかじめ用途に合わせて抽出し、データの要約などの加工をしておくこと
- イ 大規模データベースから、統計や推論の手法を使って、意味のある情報を見つけ出すこと
- ウ 大規模データベースにおけるスタースキーマを実装するのに適した、索引ファイルを作成すること
- エ 大規模データベースにおけるメタデータを効率よく管理すること

解答解説

データマイニングに関する問題である。

データマイニングは、小売店の販売データや電話の通話履歴、クレジットカードの利用履歴など、企業に大量に蓄積されるデータを解析し、その中に潜む項目間の相関関係やパターンなどを探し出す技術である。

アはデータマート、イはデータマイニング、ウはビットマップ索引、エはデータディクショナリ／ディレクトリである。求める答えはイとなる。