

データベースシステムの 基礎概念

データベースの定義

データベースという用語には多くの
解釈がある。

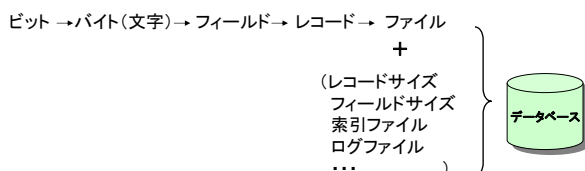
データベースの定義

- 北川博之: データベースシステム、昭晃堂 (1996).
 - 複数の応用目的での共有を意図して組織的かつ永続的に格納されたデータ群
 - →この場合にはコンピュータを使うことは暗然な条件である。

データベースの定義

- D.M.クロエンケ著、栗原潔訳: データベース処理: 基礎・設計・実装、トッパン(1996)
 - データベースとは統合レコードの自己記述的な集合である

ファイルとデータベース



データベースの定義

- 上林弥彦: 巨大データの世界、共立出版 (1994年)
 - 特定の目的に応じて大量のデータを集めたもの
 - →この定義では、コンピュータを使うことなく、住所録や名刺フォルダから企業の顧客リストまで、さまざまなものはデータベースである。

データベースの定義

- 日本工業規格 (JIS)
 - 1つ以上のファイルの集まりであって、その内容を高度に構造化することによって、検索や更新の効率化を図ったもの。

7

データベースの定義

- 通商産業省のデータベース台帳に関する規則における定義 (第2条)
 - データを整理・統合し、電子計算機による検索を行いうる形態にした集合体をいう。

8

データベースの定義

- 『データベース白書 2001』の定義
 - ① 特定のテーマに基づいて、データ体系的に整理または整理のつく状態で保存したもの。
 - ② データの集まりの中から必要なものを指定して、情報としての部分データとして取り出せるもの。
 - ③ パソコンや携帯情報端末などのコンピュータ機能を備えている情報端末機器で検索可能な形態になっているもの。

9

データベースの定義

- この講義では、この言葉を下記5つの意味が同時に成立ということで使用する。
 - (1) 特定の目的をもつこと
 - (2) データを集めること
 - (3) データを体系的に整理すること
 - (4) データを永続的に保存すること
 - (5) コンピュータを使うこと

10

なぜデータベース(システム)か

11

データの集め

- 特定の目的に応じて大量のデータを集めたもの (データの蓄積)
 - 例: 住所録や名刺フォルダ
 - 例: 企業の顧客リスト

12

データの分類

- 個人データ
 - 住所録
- 辞書データ
 - 百科事典
- 大規模データ
 - 図書情報, 新聞情報

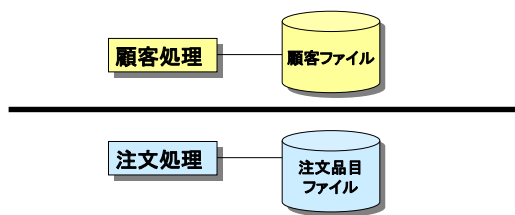
13

コンピュータを使って処理

- データの数が増えるにつれ, ことの重大性も増える.
 - コンピュータを使ってデータの管理

14

ファイルシステム



File System: オペレーティングシステムにより提供される, ファイルやディレクトリを扱うための処理機構

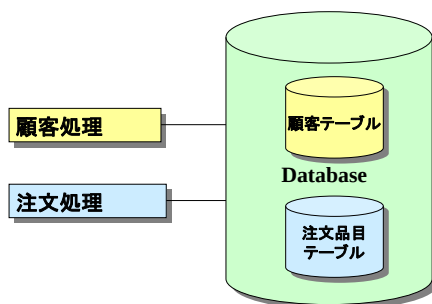
15

ファイルシステムの問題点

- データの重複が多い。
- データが分離されている。
- データがアプリケーションプログラムに依存している。
- 障害時データの保護や機密保護などが困難である。

16

データベースシステム



データ共有による一元管理

17

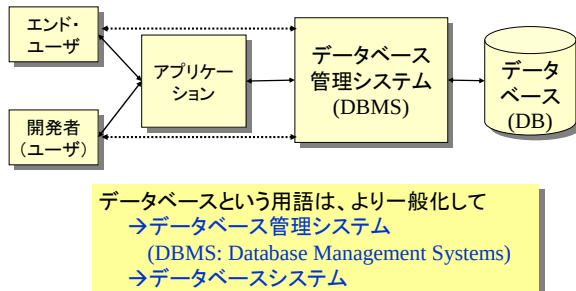
データ共有の課題

- コンピュータ複数, データは一元的に管理
- 課題
 - 複数ユーザの競合
 - 障害時回復機能
 - セキュリティ
- 例:
 - 座席予約管理システム
 - 銀行

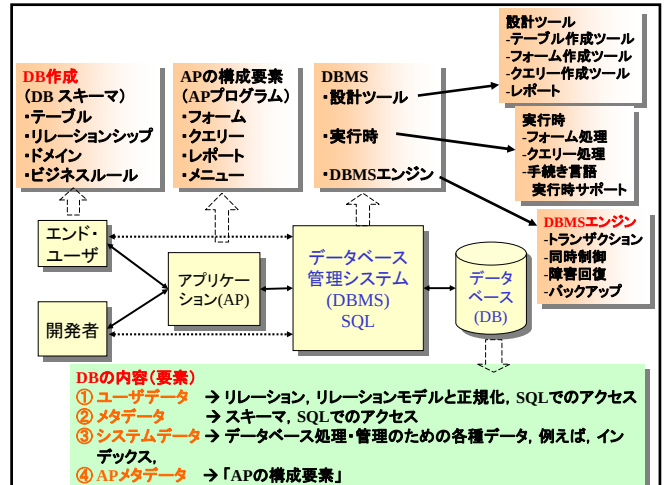
➡ DBMS: データベース管理システム

18

データベースシステムの構成要素



19



アプリケーションプログラムとDBMSの関連性

- DBMS製品の機能は大幅に強化され、DBMS自身がアプリケーションのかなりの部分を処理できるようになった。
 - ほとんどのDBMS製品はアプリケーションと統合して使用できるレポートやフォーム作成の機能をもつ。
 - 共通に利用される機能をまとめて種々のデータベースアプリケーション(データベース利用システム)を開発しやすくなる。
 - Visual dBASE, Microsoft Access

21

データ独立

- データとプログラムとを分離する独立性
- データの形式や構造が変化しても、また記憶媒体などが変わっても、アプリケーションプログラムの修正は最低限で済む

22

DBMSの主な機能

- データベースの管理機能
 - データの統合
 - メタデータの定義・操作(問い合わせ処理など)
- トランザクション処理
 - 同時実行制御
 - データの障害と回復機能
- セキュリティ機能
 - アクセス制限
 - ログイン・パスワードの概念
- ユーザ・インタフェース
 - よりユーザが利用しやすい環境

23

DBMSの持つ基本的な機能

- データ定義
 - DDL (Data Definition Language)
 - 表の枠組み(スキーマ)の定義、表の削除
- データ制御
 - DCL (Data Control Language)
 - 表の操作を行うデータベース利用者に対する制限(制御)、トランザクション管理
- データ操作
 - DML (Data Manipulation Language)
 - 表のデータの登録・修正・削除、データ検索

24

トランザクション

25

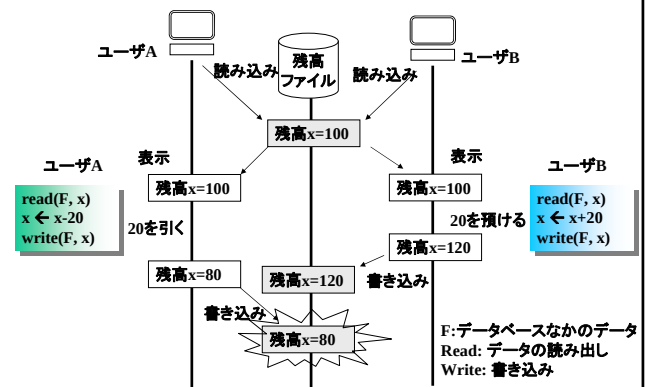
トランザクション(Jim Gray 1980)

- Jim Gray (Microsoft Research) は1998年度Turing Award 賞を受賞
<http://www.acm.org/awards/turing/citations/gray.html>
- アプリケーションにおけるひとまとまり処理を構成するデータベース操作の集まりをトランザクション(transaction)という。
- トランザクションは、分割できない処理の単位という意味合いから作業の論理単位(logical unit of work) (つまりprogram上の規定)とも呼ばれる。

26

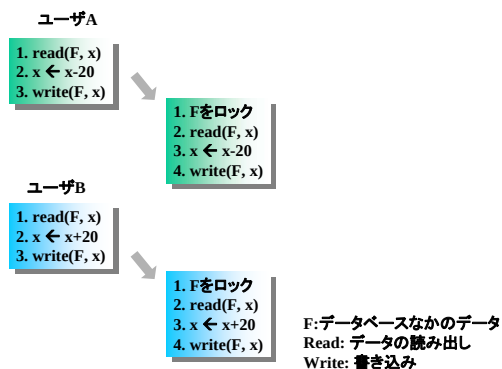


現金降ろし(同時書き込み問題)



28

ロック制御



29

ロック制御 (cont'd)



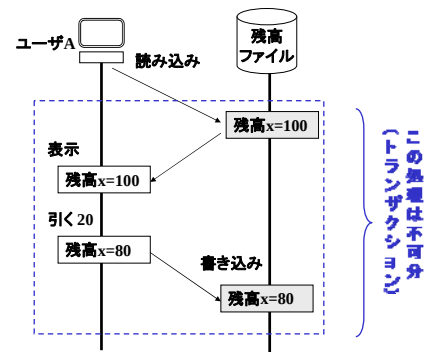
30

トランザクションの例

- 例えば、銀行で入金すること
- この入金処理には2つの作業
 - 1. お金を預かる
 - 2. 預金残高をプラスする
- この二つの作業が正常に行われた状態を入金といい、ひとつのトランザクションとなる。

31

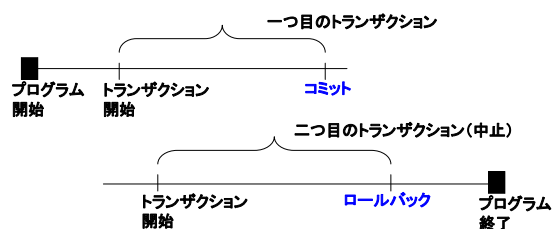
現金降ろし(正しい場合)



32

コミットとロールバック

- コミット(commit)処理
 - トランザクション終了時に、正しいデータとしてデータベースに反映させること。
- ロールバック(rollback)処理
 - トランザクション開始以前の状態に戻すこと。



33

トランザクションの4つの性質: ACID特性

- 1. 原子性 (Atomicity)
 - トランザクションは不可分である。
- 2. 整合性 (または一貫性) (Consistency)
 - 整合性がとれたデータベースに対して実行されたトランザクションの実行後のデータベースの状態は、**再び整合性**が取れるべきである。

34

トランザクションの4つの性質: ACID特性 (cont'd)

- 3. 分離性 (隔離性) (Isolation)
 - 複数のトランザクションを**並行処理**した場合でも、トランザクションは同時に処理されている他のトランザクションの影響を受けず、
 - その結果は**逐次処理**した場合と一致しなければならない。

35

トランザクションの4つの性質: ACID特性 (cont'd)

- 4. 耐久性 (持続性) (Durability)
 - トランザクションがコミットされた場合、たとえその直後にシステムの障害があっても、その変更は生き残る。

36

障害回復 ロールフォワード ロールバック

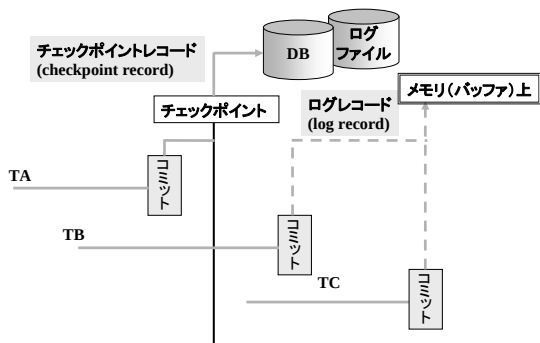
37

データベースの障害の種類

- 媒体障害
 - 例えば、ハードディスクのヘッドが壊れたなど、データの読み出しができなくなる。
 - 対策：
 - ハードディスクの二重化。
 - 定期的にバックアップをとる。
- システム障害
 - 例えば、電源が落ちた場合など。
 - 対策：
 - ログファイル(更新履歴情報)
 - チェックポイント

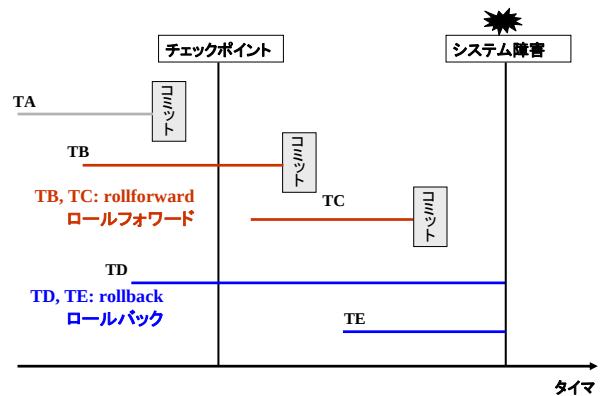
38

ログファイルとチェックポイント



39

システム障害の回復手順

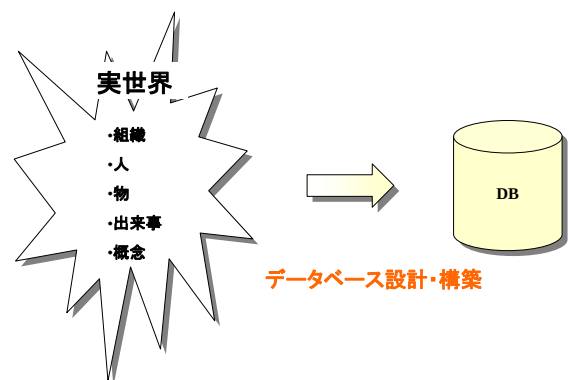


40

データベースシステムの利点

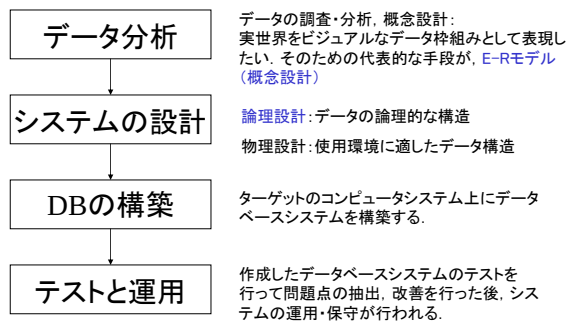
- 大量の情報の扱い
- データの多様なアクセスが可能
- データを共同利用できる
- 応用プログラムの生産性が上がる
- データの集中管理ができる
- データの更新時の矛盾を減少できる
- システムの標準化につながる

41



42

データベース設計・構築の手順



43

練習問題

- データベースシステムの構成要素と機能を説明せよ。
- DBMSの主な機能を述べよ。
- トランザクションとは？
- トランザクションの4つの性質を説明せよ。
- データベース設計・構築の手順を述べよ。

44

練習問題

- データベースという用語を定義せよ。
- ファイルとデータベースの違いについて説明せよ。
- データベースシステムでもっとも重要なものがどれか
 - 高性能ハードウェア
 - APプログラム
 - DBMS
 - 蓄積されたデータ

45