

1. データベースシステムとは

データベースシステム(database system)は、事務処理、科学技術計算、設計(CAD)、銀行等のオンライン処理、またグループウェア等の種々のコンピュータ応用の中心として用いられてきている。本章では、こうした、データベースシステムとは何か、について考える。

1.1 統合化から分散化

データベースシステムの目的は、実世界の一部(slice of reality)をコンピュータ内に実現することである。実世界とは、情報化の対象とされている世界であり、会社、部、工場、学校、個人、病院といったものである。こうした世界は、ほとんど無限の情報と、これに基づいた多種多様な活動から構成されており、そのすべてをコンピュータ内に実現することはできない。したがって、その一部がデータベースシステムとして実現される。

このことから、データベースシステムを考えるとき、まず実世界のモデルが必要となる。実世界をモデル化したものを組織体(enterpriseまたはcommunity)とする。学校、会社といった情報化の対象が組織体であり、組織体は種々の情報に基づいて種々の活動を行っている。おのこの活動を、組織体の応用または業務(application)という。応用は、具体的には、データの集合とこれに対するデータ操作を行うプログラム(応用または業務プログラム)から構成される。

例えば、学校を例として考えてみる。学校の業務として、学生の成績管理、カリキュラム管理、授業料管理、教室管理といったものがある。成績管理では、各学生とその成績のデータが必要であり、カリキュラム管理では、授業科目、先生、教室、授業時間といったデータが必要となる。これに対して、データベースシステムでは、学校という視点に立って、学校全体の業務が必要とするデータの構造を1つに定めて、これ

1. データベースシステムとは

らのデータを学校として管理していくものである。すなわち、組織体内の全応用が必要とするデータを統合化したシステムである。統合化(integration)とは、複数の応用が、それぞれの形式で必要としているデータを、組織体の目的に合うような1つの形式にまとめることである。このことは、各応用単位に管理されていたデータを、組織体としてデータを集中管理することである。

以上から、データベースシステムは、以下のように定義できる。

【データベースシステム】 データベースシステムとは、企業、学校、工場、自治体等の組織体の運用上必要となるデータを、統合的に管理するシステムである。

ここで、データベースシステムの歴史的な経緯について少し考えてみる。データベースシステムが登場してきたのは、1960年代である(図1.1)。それまでは、各応用プログラムごとに、データを保有し、管理する形態であった。プログラムとデータが一体化しており、プログラムはデータ依存していた。このようなデータの管理システムをファイルシステム(file system)といい、データの集合をファイルという。ファイルシステムとは、各応用ごとに必要となるデータの集まりをファイルとして管理するシステムである。

このファイルシステムは、コンピュータの二次記憶内のデータを利用するために、コンピュータの登場以来利用されてきたが、1950年代の後半頃から、コンピュータの利用分野の拡大に伴い、以下の点が問題となってきた。

【ファイルシステムの問題点】

- ① 複数のファイル間でのデータの一致性の保持が困難である。例えば、ある会社で、人事と保険の応用があり、おのおのの応用が社員についてのデータをファイル P と I 内に保有しているとする。このとき、ある社員“太郎”が入社したとすると、人事 P と保険 I の2つのファイルシステム内のデータを更新する必要がある。
- ② 複数のファイル内に同一のデータが存在するので、組織体全体での物理的な記憶量が増大してしまう。人事と保険の例で、各社員のデータは、人事と保険の2つのファイルシステム P と I の中に冗長に記憶されている。このために、2倍の記憶量が必要となる。

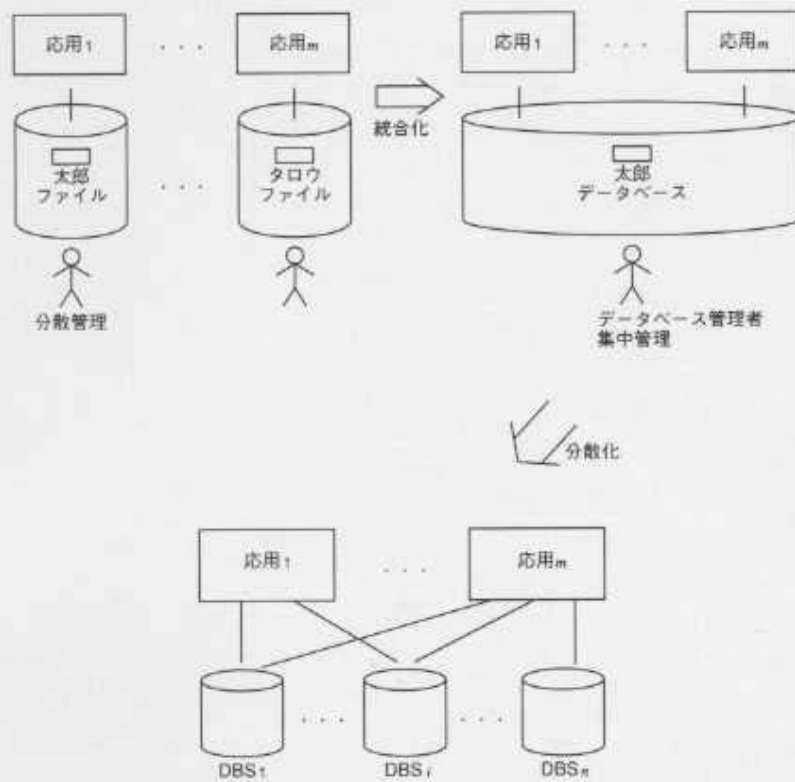


図1.1 統合化と分散化

- ③ ファイルごとにデータの表現形式が異なることにより、ファイルシステム間でのデータの共有が困難である。例えば、人事のファイルシステム P はデータをカタカナコードで記憶し、保険ファイルシステム I はローマ字で記憶していたとする。このとき、ある社員について人事と保険のデータを照合しようとしたとき、コードが異なっているために困難である。

以上のように、ファイルシステムでは、各応用ごとにおのおのが必要とするデータが分散して管理される。これに対して、データベースシステムでは、応用の集合であ

1. データベースシステムとは

る組織体ごとにデータが集中管理される。すなわち、組織体内の各要素、例えば、各社員に対して1つのデータを作り、こうしたデータの集まりをデータベースとすることである。各応用は、各自のファイルを利用する代わりに、1つのデータベースを他の応用と共有して利用することになる。

データベースシステムは、ファイルシステムに対して、以下の利点を持つ。

【データベースシステムの利点】

- ① データの冗長性を除去できる。ファイルシステムでは、組織体内のデータが複数のファイルシステム内に冗長に記憶されていたが、データベースシステムでは、1つのデータだけがデータベースに記憶される。このために、データの物理的な記憶効率を向上できる。
- ② データの一致性を保つことが容易である。同一のデータがデータベースシステム内に1つだけ記憶されているので、更新をこのデータに対して行えばよくなり、データの一致性を保つことが容易になる。
- ③ データの共有性を向上できる。データが標準の形式で記憶されるので、複数の応用によるデータの共有が容易になる。
- ④ データの安全性を向上できる。各応用が直接に物理的なファイルを操作せず、標準的なインタフェースを通してのみデータを操作させることにより、データの安全性を向上できる。
- ⑤ 組織体内の標準化を行える。組織体としての標準のデータ形式、データ操作方法を与えることができる。

このように、統合化の概念に基づいて、データベースシステムが登場してから、広く利用されてきている。60年代から70年代にかけての情報システムは、大型の汎用コンピュータを中心に構成されていた。データベースシステムも、大型汎用コンピュータ内に実装され、利用者は、端末を通じてデータベースシステムを利用していた。

1980年代から、高性能なワークステーションが情報システムの主要な要素となり、かつローカルエリアネットワーク(LAN)、広域ネットワーク(WAN)といった通信ネットワークにより相互結合されてきている。こうした情報システムでは、データベースシステムは大型コンピュータ内だけでなく、各利用者のワークステーションにも構成されてきている。例えば、大型コンピュータのデータベースシステム内のデータをワー

クステーションにダウンロードして利用する例がある。このように、統合化の概念のもとに、1つのコンピュータ内に集中化されてきたデータベースシステムが、80年代から、分散化されだしてきている(図1.2)。各応用単位に管理されてきたデータベースシステムが、情報システム全体としての緩やかな管理に従っていく形態であり、これを連合化(federation)という。

[分散化の理由]

- ① ワークステーション等のコンピュータの低価格化、高機能化、高性能化により、大型コンピュータを用いるよりも、複数のワークステーションを用いてシステムを構成する方が、安価となっている。
- ② 1つのデータベースシステム内のデータだけでなく、複数の種々のデータベースシステム内のデータを利用することが求められてきている。

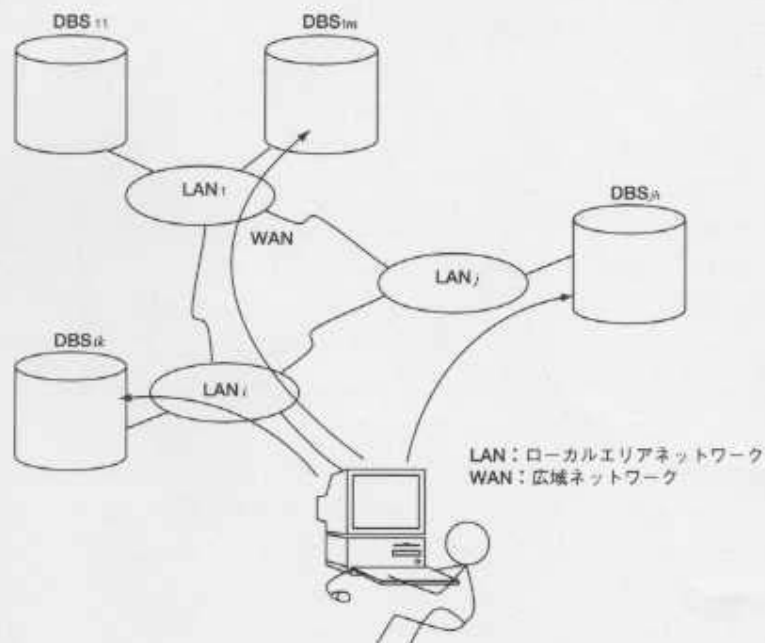


図1.2 データベースシステムの分散化