

第2章 テクニカルスキル

2.1 データベースの知識

情報システムの中でデータは最も重要な資源の一つです。データベースは、複数のデータを一元的に管理し、共有することによって検索や更新を効率良く行えるようにしたシステムです。

2.1.1 データベースの特徴

(1) データの独立性

特定の業務などに関連したデータを一つのデータ群にまとめて、データ同士を関連付けて管理することにより、データの矛盾や重複がなくなり、保守性も良くなります。

(2) データの共有・標準化

データの一元管理により整合性が保たれるとともに、標準化が促進されユーザの利用形態に応じて自由に利用できます。

(3) セキュリティ・障害対策

利用者を限定し、不正なアクセスを防ぐことでセキュリティやプライバシーが確保されるとともに、障害発生時には発生直前の状態に回復することができます。

2.1.2 データベースの種類

種 類	内 容
ツリー型	階層型データベース。親レコードと子レコードがポインタによって結合される木構造のデータベース。一つの親レコードに対し、一つまたは複数の子レコードが存在する
ネットワーク型	網型データベース。一つの子レコードが複数の親レコードに属することができるネットワーク状のデータベース
リレーショナル型	関係型データベース。1970年E.F.Codd氏によって提唱されたリレーショナルデータベース。一件のデータ（行）を複数の項目（列）の集合で表現し、データの集合をテーブル（表）で表す。データベースの操作にはSQL言語を使用する
キーバリュ型	キーによる問い合わせで値（バリュー）を取得するタイプのデータベース。高速性に特徴がある。分散処理に向いており、サーバの数を増やすことによって性能を拡張できる

2.1.3 データベース管理システム (DBMS)

共有データとしてのデータベースそのものを管理し、データに対するアクセス要求に応える専用のソフトウェアをいいます。

データベース管理システムには次のような機能があります。

- ・ データベースの定義、初期化、生成
- ・ データベースの管理 (拡張・再編成)
- ・ アクセス制御
- ・ 領域管理 (インデックス・ポインタ方式)
- ・ 機密保護 (アクセス制限・更新権限)
- ・ 排他制御 (ロック・アンロック機能)
- ・ 障害回復 (ログ機能・チェックポイント・リスタート機能)

2.1.4 リレーショナル型DBMS (RDBMS)

(1) データモデル

リレーショナルデータベースでは、人や物などのデータの集まりの関係を三つのデータモデルとして定義します。

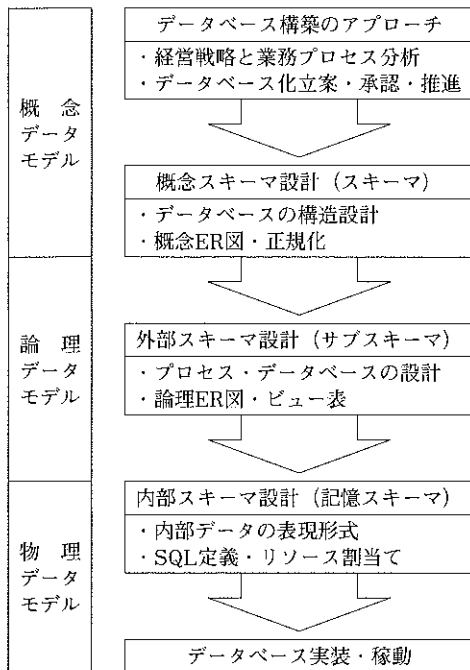
種 類	内 容
概念データモデル	データベース全体を表現したモデル。ERモデルは概念データモデルを図式化したもの
論理データモデル	利用者やプログラムから見たモデル
物理データモデル	コンピュータ内部の格納形式を表現したモデル

(2) 層スキーマ構造

データベースの定義情報をスキーマといい、論理構造や物理構造を定義します。三つの独立した層で構成されています。

種 類	内 容
概念スキーマ	データベースの「設計者」に対応し、データベース全体 (実表) を定義する
外部スキーマ	データベースの「利用者」に対応し、ビュー表など利用者やプログラムから見た定義を行う
内部スキーマ	データベースの「構築者」に対応し、コンピュータ上に実装するための定義を行う

2.1.5 データベース設計の流れ



2.1.6 データの正規化

データベースは、複数のデータ項目から構成されています。これらのデータ項目を関連付け、一つにまとめた情報としてグループ化する手法を正規化といい、一事実・一箇所の状態を構築して行きます。

正規化	内 容
第1正規化	レコード内の繰り返しの部分を排除し、独立したレコードとして分割する (別行か別表にする)
第2正規化	キーとなる項目が決まれば特定のデータが決まるように、キーと従属関係のない項目を排除し、別表として分割する
第3正規化	キー以外の項目同士での従属関係 (推移従属) を排除し、別の表として分割する

2.1.7 多様化するデータベース

(1) 分散データベース

ネットワーク上に物理データベースを分散しているデータベース。利用者は、どこにあるか意識しないで利用できるが回線速度・コスト・障害時の局所化・互換性などを検討して設計する必要があります。

(2) マルチメディアデータベース

数値や文字以外に、図形・画像や音声などのデータを格納や利用を行うものでオブジェクト指向の形態に近いものがあります。

(3) オブジェクト指向データベース

文字データ・画像や音声などのデータの集合を手続き（操作）を含めて一体化したオブジェクトとして扱うデータベース。

(4) ハイパーテキストデータベース

テキストとテキスト（絵や表）のリンク（関連）を設定してオブジェクトをたどっていけるデータベース。自由度が高く、インターネットのホームページ検索などに利用されます。

(5) パーソナルデータベース

主に個人で利用するように作られたデータベースでデータベースの知識が無くても構築や利用できる。カード型やリレーショナル等の簡易型から本格的なデータベースまであります。

高度なデータベースの利用例

(1) データウェアハウス

履歴型データを大容量のデータベースに蓄える倉庫であり、意志決定支援や多次元データ分析等に利用されます。

(2) データマイニング

膨大な量の詳細な生データ（データベース）から、マーケティング情報などの分析で一定のルールや仮説を導き出して活用します。

2.1.8 SQL (Structured Query Language)

リレーショナルデータベースを操作するために開発された言語で、「構造化問合せ言語」という意味付けがなされ、その頭文字をとって「SQL」と名付けられました。

SQLは規格化されたことにより互換性が高くどのRDBMSでも定義や操作が可能になり利便性が向上しました。

2.1.9 SQLの進化

(1) SQLの進化

SQL	内 容
1 SQL2	1992～1995年。1992年にISO、JISによるSQL92の規格化。DDLでは参照制約動作やデータ型の拡張など、DMLでは動的機能、関係演算、トランザクション機能などが拡張されている
2 SQL3	オブジェクト指向や、マルチメディア関連の拡張を行った規格であり、1999年に世界標準規格として制定
3 SQL4	SQL国際標準委員会は、次のSQL標準（開発コード名は「SQL4」）として次課題を検討している

(2) SQLの構成

SQL文の対話と関数で、データベースの操作を行います。

操 作	内 容
対話	<ul style="list-style-type: none"> ・ CREATE SCHEMA スキーマ定義 ・ CREATE TABLE 実表定義 ・ CREATE VIEW ビュー表定義 ・ GRANT 権限定義
操作	<ul style="list-style-type: none"> ・ SELECT 検索 ・ INSERT 挿入 ・ UPDATE 更新 ・ DELETE 削除
関数操作	<ul style="list-style-type: none"> ・ SELECT INTO 検索 ・ INSERT 挿入 ・ UPDATE 更新 ・ DELETE 削除
	<ul style="list-style-type: none"> ・ DECLARE 宣言 ・ OPEN 開く ・ FETCH 取出し ・ CLOSE 閉じる
	<ul style="list-style-type: none"> ・ COMMIT 確定 ・ ROLLBACK 取消

2.1.10 NoSQL (Not Only SQL)

リレーショナルデータベース (RDBMS) は柔軟性に弱点があります。不特定かつ大量のリクエストに対する応答や膨大なデータの扱いが求められる場合には、性能限界を考慮しなければなりません。

この弱点を克服するため、単純なデータ構造を持つ膨大なデータに対して瞬時に検索するデータベースとしてNoSQLと呼ばれる種類のデータベースが存在します。NoSQLは、RDBMSの持つ高度な機能を排除することで、シンプルで高い柔軟性を獲得しています。

項目	NoSQL	RDBMS
メリット	大量データに対して高速検索が可能 データが分散されるため単一故障点がなく耐障害性が高い 性能向上が容易 規模の経済性を享受しやすい	強力なトランザクション機構による高度な一貫性の保証 SQL言語による汎用的な問い合わせが可能
デメリット	複雑な問い合わせが出来ない 一貫性、可用性のいずれかを犠牲 問い合わせ方法がデータベース毎に異なる	性能向上が困難 インデックス作成にコストがかかる
データモデル	データベース毎に異なる ・カラムグループ型 ・キーバリュ型 ・ドキュメント型 ・グラフ型	リレーショナルモデル
問い合わせ	キーによる高速問い合わせ	SQLによる高度な問い合わせ
性能向上策	サーバの並列配置	高速なハードウェアの導入 データの水平分割配置

データベースは、一貫性 (Consistency)、可用性 (Availability)、分断耐性 (Partition Tolerance) の全てを同時に満たすことは出来ません。RDBMSは一貫性を重視しているのに対し、NoSQLでは、一貫性か可用性のいずれかを犠牲にして、分断耐性を最重要視していることが大きな特徴のひとつです。

2.2 ネットワークの知識

ネットワークとは、一般的にデータを伝送する通信網のことです。コンピュータにおけるネットワークは、複数のコンピュータを接続して、データやプリンタを共有し、情報や資源を効率よく使う技術です。

また、モデムやルータを使いNTTなどの通信業者の回線（電話、ISDN、専用線など）を利用した遠隔地との広域ネットワークを世界規模に発展したものがインターネットです。

2.2.1 TCP/IP

データ転送を行うための通信規約をプロトコルと言います。送信側と受信側は同じプロトコルを使います。

TCP/IPはインターネット、イントラネットにおける通信プロトコルの総称で「TCP」と「IP」という二つの代表的プロトコルを合わせたものになります。

TCP/IPによる通信はインターネットプロトコルとして急速発展し、現在では世界標準の通信プロトコルになっており、インターネットへのアクセスが可能なコンピュータやパソコンは、全てTCP/IPを利用しています。

プロトコルの標準的なモデルに「OSI参照モデル」があります。「OSI参照モデル」はITU（国際電気通信連合）やISO（国際標準機構）などが、異なるコンピュータ間でも接続や通信を可能とするネットワーク構築に必要な全体構成や各々のプロトコルなどの標準化を進めて出来たものです。

OSI参照モデルでは、プロトコルをより細かく、7つの階層に区別しています。これにTCP/IPプロトコルを照らし合わせると、TCPはトランスポート層にあたり、セッションという形で1対1の通信を実現し、エラー訂正機能などを持ちます。

IPはネットワーク層にあたり、パケットを相手に届ける役割を持ち、発信者、受信者（手紙でいう宛て先）、利用経路などの情報を含めることができます。発信者、受信者は、IPアドレスにより特定します。発信者から送信されたパケットは配送先と分岐先の対応表であるルーティングテーブルに従ってルータにより必要に応じて別のネットワークへと中継され、受信者へと届けることができます。

IPは、相手にパケットを届けることのみを目的としており、相手まで確実にパケットが届いていることを保証しないので、途中でパケットが失われた場合には届きません。確実な送信をするには上位プロトコルのTCPを利用します。

現在、主に利用されているのは32ビットのアドレス空間を持つIPv4であり、IPアドレスの不足が発生することから128ビットのアドレス空間を持つIPv6に移行されています。

階層	OSI参照モデル	機能概要	プロトコル
第7階層	アプリケーション層	メールやファイルの送受信等データの具体的な利用方法	HTTP、SMTP、POP、FTP、TELNET、NNTP、DNS
第6階層	プレゼンテーション層	文字のコードや画像データの形式等、データの表現形式	
第5階層	セッション層	通信の開始と終了	
第4階層	トランスポート層	誤りがあった場合の手順、パケットの抜けの確認手順等	TCP
第3階層	ネットワーク層	接続するコンピュータのアドレスに関する取り決め	IP
第2階層	データリンク層	送受信するパケットの区切りや構成についての取り決め	
第1階層	物理層	ケーブルやコネクタ、電圧、電気信号の速度などの取り決め	

2.2.2 インターネット

インターネットは「世界最大のコンピュータネットワーク」とよく表現されます。

具体的には、LANと呼ばれる企業や大学や研究機関などのネットワークを相互に接続したネットワークです。米国防総省が始めたコンピュータネットワークを起源とし、その後、徐々に世界規模に発展してきたものです。

インターネットは、TCP/IPという機種に依存しない通信プロトコルを利用しているため、インターネットに接続したコンピュータは、機種の違いを超えて通信することができます。また、インターネットは、パソコン通信のように1つのホストコン

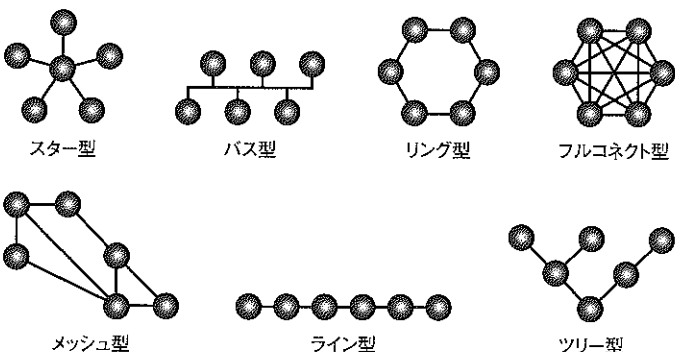
第2章 テクニカルスキル

コンピュータがすべてのサービスを提供しているわけではなく、全世界に無数に存在するサーバコンピュータが、相互に接続してサービスを提供する分散型のネットワークとなっています。

インターネットはブロードバンド化と、モバイル化が進展し、映像、音楽などの大容量データを扱う新しいサービスが生まれ、インターネットの利用用途は拡大しています。

2.2.3 ネットワークトポロジー

ネットワークトポロジーとは、物理的なネットワークの接続形態を示し、各端末や制御機器がどのような形態で接続されているかを現します。代表的なトポロジーには「スター型」「バス型」「リング型」「フルコネク外型」



2.2.4 無線LAN

ケーブルによる配線を使わないLANのことで、物理層レベルで分けて電波を使うものと赤外線やレーザーを使うものがあります。

無線LANのメリットとしては、

- ・ケーブルが要らないので配線スペースが不要
- ・末端の設置や移動が自由
- ・移動体での使用が可能
- ・迅速なLANの構築が可能
- ・屋外通信が可能

などがあげられます。

無線LANの規格についてはIEEEの802.11のワーキンググループで決められています。

(IEEE:米国電気電子学会)

2.2.5 ネットワークセキュリティ

ネットワークセキュリティとは、ネットワーク上での安全確保のための防衛策で、システム攻撃者からコンピュータシステムを守り、不正アクセス防止や情報漏えいの阻止、システムの安定性保持を行うことです。

インターネットの普及により、外部からのシステム攻撃者の脅威が取りざたされるようになったため、外部ネットワークとの境界にファイアーウォールを設置して内外のデータの流通を監視し、不要なアクセスを制限するのが一般的になっています。

(1) ファイアーウォール

ファイアーウォールとは、「信頼できるネットワーク」と「信頼できないネットワーク」の2つのネットワーク間のアクセスを制御するために使われます。

組織内のネットワークでは、インターネットなどの外部ネットワークを通じて第三者が侵入し、データやプログラムの盗み見・改ざん・破壊などが行われることのないように、外部との境界を流れるデータを監視し、不正アクセスを検出・遮断するシステムです。

防御方法	説明
パケット フィルタリング型	送信者や送信先のIPアドレス、ポート番号などによって通信データを通過させるかどうかを判断し、不正アクセスを防ぐことができます。OSI参照モデルでいうとネットワーク層で動作します
アプリケーション ゲートウェイ型	通信を中継するプロキシサーバを利用し、組織内のネットワークと外部のネットワークの間で直接通信を出来ないようにする方式。OSI参照モデルでいうとアプリケーション層で動作します

(2) 暗号化

暗号化とは、データを第三者には分かりにくい形のデータに変換することです。暗号化の逆の操作を復号化といいます

データを暗号化するには、データを変換するための規則（アルゴリズム）と、鍵（キー）2つを使います。暗号化／復号化のためのアルゴリズムをデータの送信者と受信者の間で取り決めて、さらにデータを暗号化／復号化するとき鍵となるデータを使ってデータを変換します。アルゴリズムと鍵の両方を使うことにより、たとえばアルゴリズムが漏えいしても鍵の値が分からない限り、データを復号化することは困難となります。

データ暗号化の方式には、大きく分けて、公開鍵と秘密鍵の対になる2つの鍵を使う公開鍵暗号方式、暗号化も復号化も同じ鍵を用いる共通鍵暗号方式があります。

	公開鍵暗号方式	共通鍵暗号方式
鍵の管理	相手が複数でも、秘密鍵は1つなので管理が容易である	相手が複数の場合、複数の秘密の鍵が必要なので複製が困難である
鍵の交換	公開鍵のみの交換で済み安全である	秘密鍵の安全な交換が必要である
鍵の交換時の危険性	改ざんに注意が必要である	盗聴されれば複製が可能である
処理時間	長い	短い
認証	第三者に認証できる	不十分である