

第8回 Web サーバとデータベース、ネットショップの仕組み

これまではブラウザの設定や知っておきたい機能を説明しました。今回はブラウザの相手側、Web サーバに目を向けてみましょう。Web サーバの仕事をつきつめると、ブラウザからのリクエストに応じて用意してある HTML ファイルを送信すること。では Amazon.com のように 1 千万点を超える商品を扱うネットショップでは、商品と同じ数の HTML ファイルが用意されているのでしょうか？そこに、最近の Web サイトの仕組みの特徴があります。今回は、Web サーバとデータベースのお話です。



【今回登場するキーワード】

「Web サーバ、アプリケーションサーバ、データベースサーバ」

「データベース」

「DBMS (Database Management System、データベースマネジメントシステム)」

「SQL (Structured Query Language、構造化問い合わせ言語)」

「CMS (Contents Management System、コンテンツ管理システム)」

「ショッピングカート」

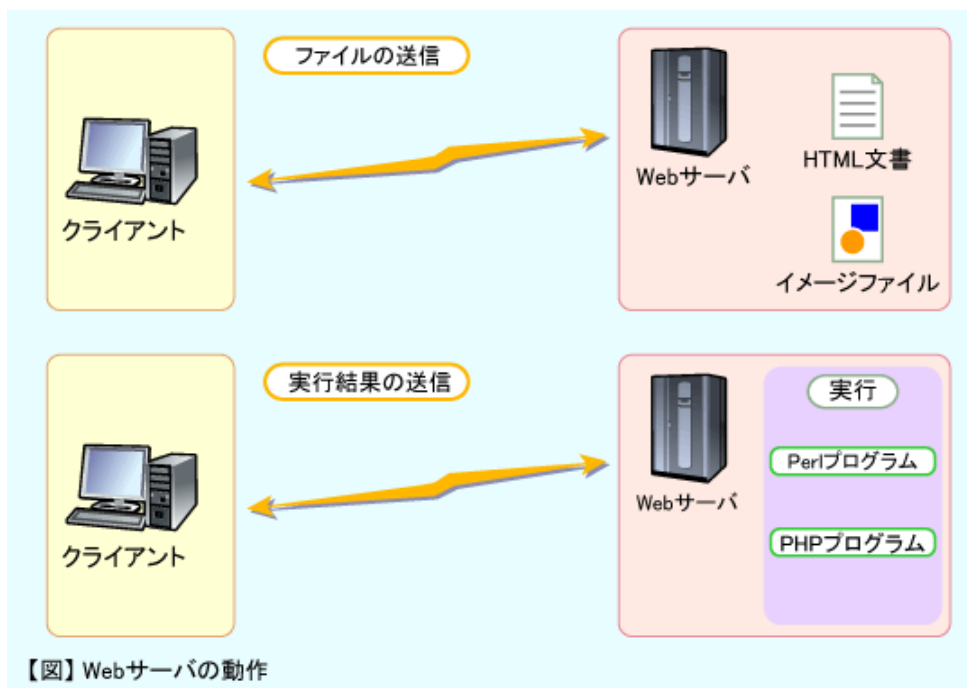
Web サーバは、ブラウザからの要求に応じて、内部に保存してある HTML 文書や画像などの情報を、インターネットなどのネットワークを通じて送信するのが役割です。しかし、最近の Web サーバは単にサーバ内に保存してあるファイルを送信するだけではすまなくなっています。プログラムを実行して結果を送信したり、結果に応じて処理を変更したり、データベースと連携して複数の処理をまとめて実行するなど、Web に求められるものが高度になるにつれて Web サーバの仕事は多様化し複雑になっています。そのため、最近の Web サイトは複数のサーバとデータベースで構成されるようになってきています。

■Web サイトとデータベース

Web サーバ、アプリケーションサーバ、データベースサーバとデータベースというのが、Web サービスなどを提供する Web の典型的な構成です。一つひとつのサーバから解説していきましょう。

・サイトの玄関口、Web サーバ

Web サーバとは、ブラウザからの要求を受け取り、要求された HTML ファイルを返す働きをするサーバソフトウェアを搭載したコンピュータのことを言います。もう少し詳しく言うと、「ブラウザからの要求を受け、HTML ファイルと画像などの付随するファイルを送信する」、「ブラウザからの要求に応じてプログラムを起動し実行結果を送信する」の 2 つの機能を果たします。



Webサーバには数十から、大型のサイトでは数千のブラウザが同時にアクセスします。それでも、サーバ内部のHTMLファイルを要求通りに送信するだけであれば、それほど大変ではなかったのです。ところが、今日のWebサイトの機能が高度化するにつれ、Webサーバの処理の内容が多様に、しかも複雑になっています。次の表はそのWebサーバの機能一覧です。

機能	内容
待ち受け	ブラウザから接続を待ち受けます。
要求の受信	ブラウザからのファイルまたは処理の要求を受け付けます。
セッション管理	ブラウザとサーバ間の通信を管理します。Cookieの生成や送信、ブラウザの識別、SSLによる暗号化通信を行います。
HTML等の送信	ブラウザへHTMLと付随するファイルを送信します。ブラウザからの処理要求に対しては、処理結果をHTMLに変換して送信します。
外部プログラムの起動	PERLやVB、JAVA、C言語などで作られたCGI(Common Gateway Interface)により掲示板やチャットなどのWebサイト上のサービスを実現したり、計算やデータ加工、HTMLの生成、データベースへのアクセスを行ったりしています。
外部サーバとの連携	ユーザー登録の確認やワンタイムパスワードなどをメールサーバと連携してメールで送信したり、企業の基幹システム等と連携して業務用アプリケーションとデータの受け渡しを行います。
アプリケーションの実行	サーバソフトウェア上で動作するサーバサイドスクリプト(プログラムの1種)を実行します。CGIと比べてサーバのCPUやメモリへの負荷が低いのが特長です。ASP(Active Server Pages、マイクロソフト)、JSP(JAVA Server Pages、IBMとサンマイクロシステムズ)、PHP(PHP、Hypertext Processor、Rasmus Leordor氏)が有名です。
データベースへの接続	サーバサイドスクリプトなどを通じてデータベースに接続し、ブラウザから要求されたデータ処理を行います。

表：Webサーバの機能

上記のような機能に加えて、Web サーバには、動作の安定性や24時間365日動作する耐久性、システム内外の障害への耐性や障害からの迅速な復帰、不正なアクセスや攻撃からデータやシステムを守る高いセキュリティ性能などが求められます。

現在、Web サーバソフトウェアとして市場を二分しているのが Apache と IIS です。Apache は、NCSA httpd Ver.1.3 をベースにオープンソースソフトウェアとして機能を拡張してきた Web サーバソフトウェアです。最初の公式リリース(0.6.2)は、1995年8月。FreeBSD、Linux などの UNIX 系 OS はもちろん、Windows (Win32 版) などにも移植されており、現在も世界中のボランティアの手で開発が進められています。プロキシ機能や SSL 対応などモジュールによる機能追加ができることとオープンソースソフトウェアであることが特長です。

Microsoft IIS は、Windows NT Server、Windows 2000 Server、Windows Server2003 に標準で付属する Microsoft 社のサーバソフトウェアです。ASP(Active Server Pages)や ActiveX による動的なページを構築するための機能や、Gopher や FTP サーバ機能も含まれます。1996年2月より無料でダウンロードできるようになりました。IIS 4.0 までは「Internet Information Server」ですが、Windows 2000 搭載の IIS 5.0 からは「Internet Information Services」となっています。

内部に蓄えた HTML ファイルを表示する限りにおいて差はありません。しかし、他システムとの連携や Web サイト、イントラネットサイト構築などの容易さでは IIS に分があるようです。

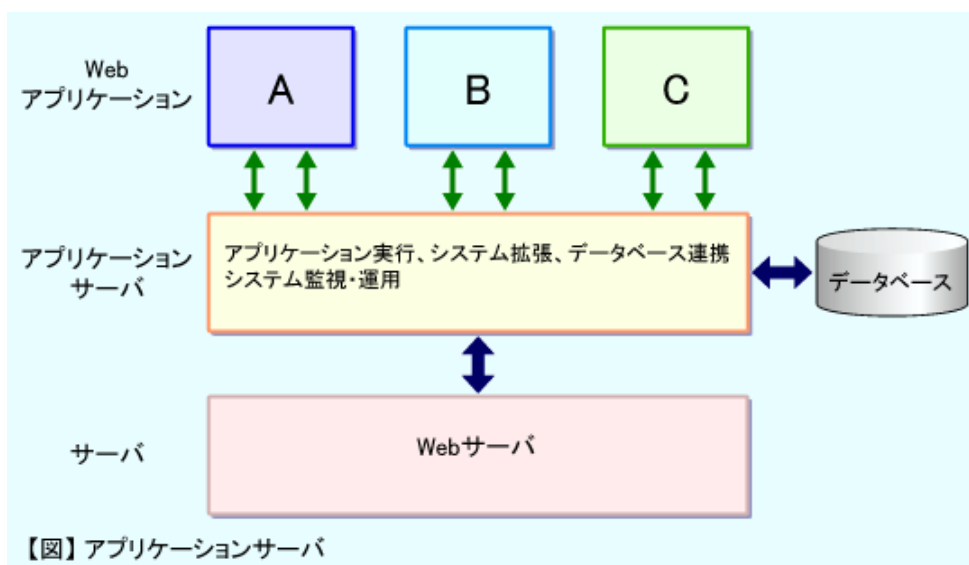
Windows との親和性が高く、企業の基幹系システムとの連携にも標準的な GUI のツールが用意されており比較的容易にサーバが構築できます。これに対し Linux+Apache にも多くのツールが出回っていますが GUI を実現したものは少なく、サイト構築にはある程度の知識やスキルが要求されます。Web サーバの 80%が Apache といわれた時期がありましたが、2007年秋には両者のシェアはほぼ互角となっています。現在、企業ベースでは IIS の利用の方が多いと言われています。

・忙しすぎる Web サーバ

掲示板やチャット、会員登録、ページのカスタマイズなどユーザーとの双方向のやり取りが発生すると、Web サーバは CGI などのプログラムを起動して、処理結果を返したり、処理結果に応じて表示を変えたり、他のサーバと連携して処理を行います。また、処理によって表示が変化するデータや大量のデータを扱う場合には、あらかじめすべての表示用 HTML ファイルをすべて用意しておくことは不可能です。通常は、そうした場合にデータベースと連携してその都度表示用ファイルを生成します。そうすると Web サーバは多忙になります。ブラウザとのやり取り、Cookie の送受信、ブラウザごとにどのページを閲覧しているかというセッションの管理などの窓口業務、データ検索や抽出を行うアプリケーションの起動、データベースへの接続、結果を表示するためにブラウザに渡す HTML の生成などの処理業務、これらを同時に複数のブラウザに対して行うこととなります。Web サーバ単体では限界があり、動的コンテンツを扱わない部分も含めた Web サイト全体のパフォーマンスが低下することになります。Web サーバがボトルネックになってしまうのです。ブラウザから見ると反応が遅い Web サイトということになってしまいます。

・ Web サーバの業務代行をするサーバ

Web サイトで発生する計算やデータの検索、抽出など、さまざまなアプリケーションを動作させるために専門のサーバを用意して、Web サーバの仕事量を減らす仕組みがあります。これをアプリケーションサーバといいます。プログラムの実行環境やデータベースへの接続、複数の処理を連結するトランザクション管理機能、業務の処理の流れを制御する機能などを実装しています。CGI や PHP などのサーバサイドプログラムを実行したり、データベースから必要なデータを取得しそのデータを元に HTML を生成したり、Cookie などのセッション管理やメールサーバと連携したり、大規模なシステムでは複数のサーバマシンを動作させる時の負荷分散処理や企業の基幹システムとの連携も行います。まさに八面六臂の活躍で、Web サーバのほとんどの仕事を肩代わりします。



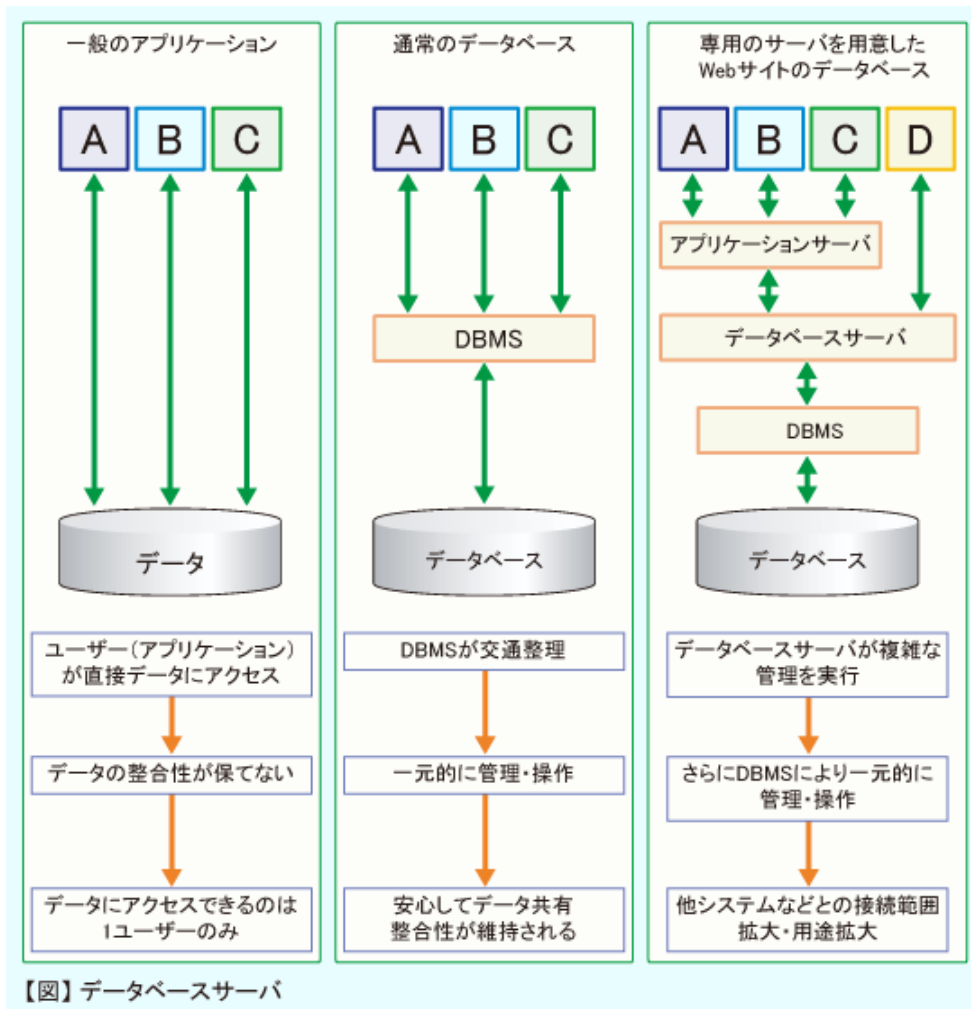
その結果、Web サーバはブラウザとの窓口業務に専念できるようになります。多くのユーザーを同時に相手にすることができ、Web サイト全体のレスポンスが向上します。代表的なアプリケーションサーバには、WebLogic、Tomcat、.NET Framework、WebSphere、Oracle Internet Application Server、ColdFusion、などがあります。

・ データベース専門のサーバ

Web サイトには、データベースサーバというデータベースの操作と管理を行う専用のサーバを設けることがあります。検索や更新などのデータベースに対する要求を受けたときに、処理を行って結果をアプリケーションに返します。

詳しくは後述しますが、データベースはデータを蓄えているデータベース本体とデータベースの操作、制御、管理を行う DBMS (Database Management System、データベースマネジメントシステム) から成り立っています。ユーザーが、データを参照したり、検索、更新、追加、削除などの操作を行ったりする場合にはデータベースに対して直接操作を行うわけではありません。データベースに対するすべての操作は、DBMS を通じて行うようになっています。複数のユーザーが同時にデータベースを利用しても競合を避けることができるように、更新中のデータを他者

が編集できないように制御したり、編集結果を速やかに反映したりすることができるのは、DBMSの一元管理によるものです。

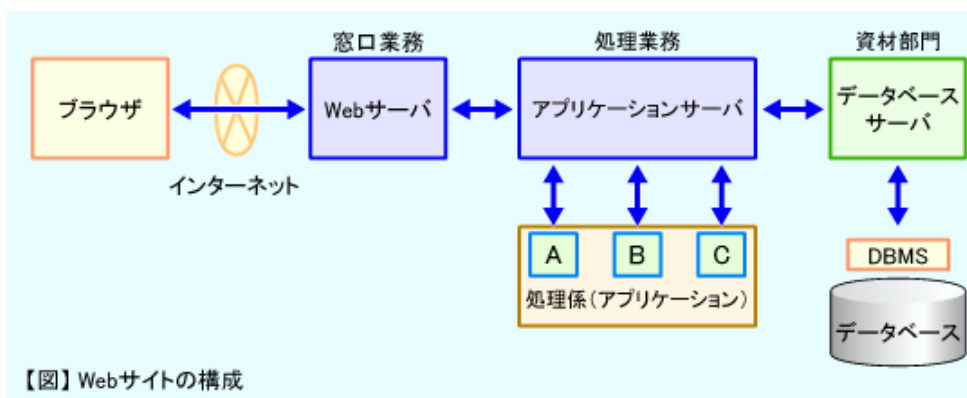


この DBMS を他のコンピュータから操作できるようにするのがデータベースサーバです。データベースの管理を専門にすることで別のシステムからのデータベース利用やセキュリティ強化、複数のデータベースの制御など、より高度な管理機能を実現します。

代表的なデータベースサーバに Oracle、SQL Server、Adaptive Server などがあります。

・ Web サイトの三つのサーバ連携の仕組み

前項で説明した三つのサーバが連携する Web サイトの構成図です。



Web サーバはブラウザから要求を受けると、アプリケーションサーバに処理を依頼します。アプリケーションサーバはプログラムを起動して処理を行います。さらに、必要に応じてデータベースサーバに接続してデータの取得を要求します。データベースサーバは DBMS に問い合わせたデータベースから結果を得て、その結果をアプリケーションに渡します。結果はブラウザ表示のための HTML ファイルに変換されて Web サーバに送られます。

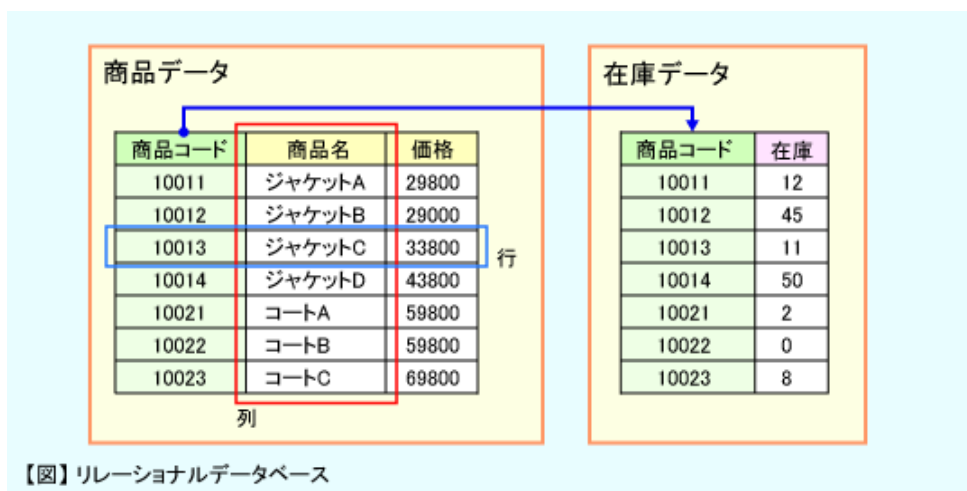
上図のように、Web サーバは窓口業務、アプリケーションサーバは計算や変換処理などアプリケーションの実行や外部のサーバとの連絡、データベースへの接続、複数の処理をひとかたまりにまとめて実行するトランザクション処理など、Web サイトの処理業務の中核を担っています。データベースはアプリケーションサーバが必要とするデータを蓄積してある資材部門、データベースサーバはその管理係です。

このような Web サイトの構成を 3 層構造システムと呼びます。ブラウザと Web サーバ間の窓口業務をフロントエンド、データベースや処理係になる業務用アプリケーションをバックエンド、アプリケーションサーバをミドルウェアと呼びます。各サーバが仕事を分担することで総合的なパフォーマンスを向上させるだけでなく、各機能が独立しているため、機能や容量の拡張をサーバ単位で行うことができます。アプリケーションの追加や削除もサイト内の他のサーバへの影響が小さく、改良や変更に強い柔軟なシステムになります。

■データベースとは情報をまとめておいて便利に使う仕組み

データベースは「情報を統合・整理・蓄積し、便利に使う仕組み」ということができます。関連する情報まで含めてすべての情報を一つにまとめ、情報に唯一無二の番号や記号を付けて蓄積します。こうして蓄積された情報を一元的に管理することで、データがどこにあるか、どのように配置されているかを意識することなく利用できるようになっています。

データベースの中で最も多く使われているリレーショナルデータベースを例にデータベースのデータの仕組みを見てみましょう。リレーショナルデータベースでは、データを次の図のように行と列からなる表形式で格納します。一つのデータベースの中にはこうした表が複数格納され、表と表は図のように関連づけられています。複数の表の列を商品番号などで結びつけて、組み合わせで在庫金額の一覧を作成したりすることができます。



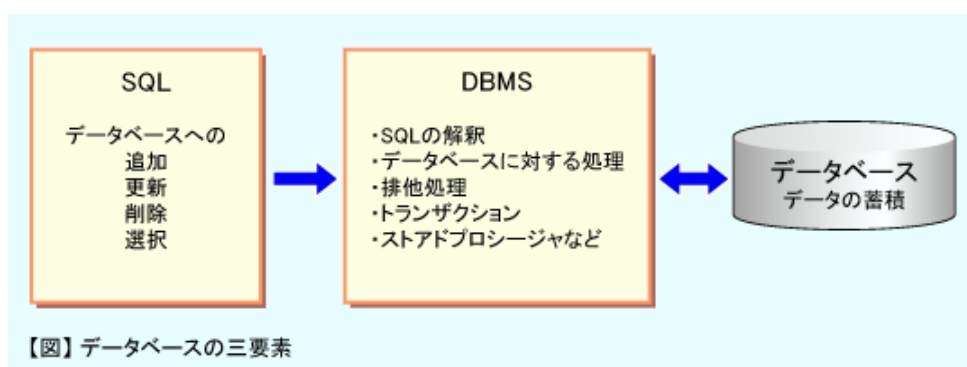
ネットショップの場合、データベースに納められた商品情報、在庫数、仕入先、注文書、発送情報などさまざまな情報を基に、Web サイトの商品カタログやおすすめ商品表示、顧客管理や売上管理などの管理情報、売れ筋情報やブログへの商品紹介などのマーケティング情報が生成されます。データが一元的にしかも一つに統合されているデータベースだからこそ可能なのです。

また、データベースは複数のユーザーやアプリケーションからのアクセスを可能にする強力な共有機能や大切なデータを守るためのセキュリティ機能など、データ保護、活用のための機能が実装されています。

機能	内容
データの独立性	データの構造が変わってもプログラムからは独立しているので、プログラムの修正が必要ない。また、どのプログラムからも利用できる。
データの一貫性	複数のユーザーが同時にデータを操作できること。また、同時に変更や削除した場合でも、データに矛盾が発生することなく利用できる。
データの整合性	データの重複が発生しない。
データの機密性	データベースへアクセスするユーザーを制御できる。
データの障害対策	データベースに何らかの障害がおこった場合でも、回復するための手段を有する。

表：データベースの機能

データベースは、「データベース」、「DBMS」、「SQL」の三つの要素で構成されます。



【図】データベースの三要素

データベース本体はデータを格納し蓄積します。データの操作はすべて DBMS が行います。データベースは情報の倉庫です。リレーショナルデータベースでは前述の図のように表の形でデータがそこに格納されています。格納されたデータの追加・更新・削除、新しい表の作成などは、DBMS の仕事です。

DBMS はデータベースの管理人です。アプリケーションとデータの間で位置して、処理要求の受け付け、処理の実行、結果の出力まで、データベースに対する処理の一切を行います。データベースとアプリケーションのやりとりはすべて DBMS を通ります。DBMS がデータベースの操作と管理のすべてを一元的に行うことで、複数のユーザーが同じデータに同時にアクセスした場合に、編集中のデータを別のユーザーが削除したり編集してしまったりしないように制御します。これを「排他制御」といい、複数のユーザーやプロセスが同時に使用することが前提となっているデータベースにとって重要な機能です。さらに、データベースの中核的な機能として、「トランザクション処理」、「ストアドプロシージャ」という機能があります。

トランザクションとはひとかたまりの処理のことをいいます。例えば、銀行口座からの振込処理では、口座から振込金額を引き出し、その金額を送金処理します。万一、口座から引き出した直後にサーバが停止すると、お金が引き出されたまま、つまり残高はマイナスされたままになります。一方で振込はできていないので、その分の金額がシステムの中で宙に浮いてしまいます。そうならないように、口座からの預金の引き出しから送金処理が終わるまでをひとかたまりの処

理とし、それが終わるまではデータを反映させないような制御が必要になります。これをトランザクション処理と言い、データ全体の整合性をとるための重要な機能です。

ストアプロシージャは、データを処理するための一連の手続き（Procedure、プロシージャ）をデータベース側に格納することを言います。ストアプロシージャを指定するだけで処理を行うことができるため、ネットワークのトラフィックを減らすことができますし、あらかじめ手順が準備されているために DBMS の負担も小さくなります。

DBMS に処理を依頼するための専用の言語が SQL (Structured Query Language、エスキューエル、構造化問合せ言語) です。この SQL によりデータベースに依頼することで、データの定義や、参照・挿入・更新・削除・表の結合などデータ操作、更新情報の整合性を保つためのデータ制御などが実行できます。SQL は DBMS が理解する唯一の言語です。言い換えると、SQL で DBMS に要求できればどのようなアプリケーションでもデータベースを利用できるということになります。言語としての SQL は「商品番号〇〇の商品名を取得する」を「SELECT 商品名 FROM 商品番号」というように、C 言語や Java のようなプログラミング言語よりも直感的に理解しやすい文法になっていますが、ユーザーが SQL をキーボードから直接入力することはありません。通常はアプリケーションが SQL で出力します。例えば、ブラウザで、ネットショップの商品を選ぶ操作をした時に、アプリケーションがデータベースサーバに対して商品の写真や在庫状況を問い合わせる SQL を送るようになっています。

なお、SQL を「シーケル」と発音する人もいます。SQL は、SEQUEL (Structured English Query Language、構造化英文問合せ言語) が基になっているからです。SEQUEL は 1970 年に登場し、1976 年に SEQUEL 2 が発表されています。しかし、1987 年に規格化されたときには、SQL87 となっており、そのころから SQL という呼び方が一般的になっているようです。

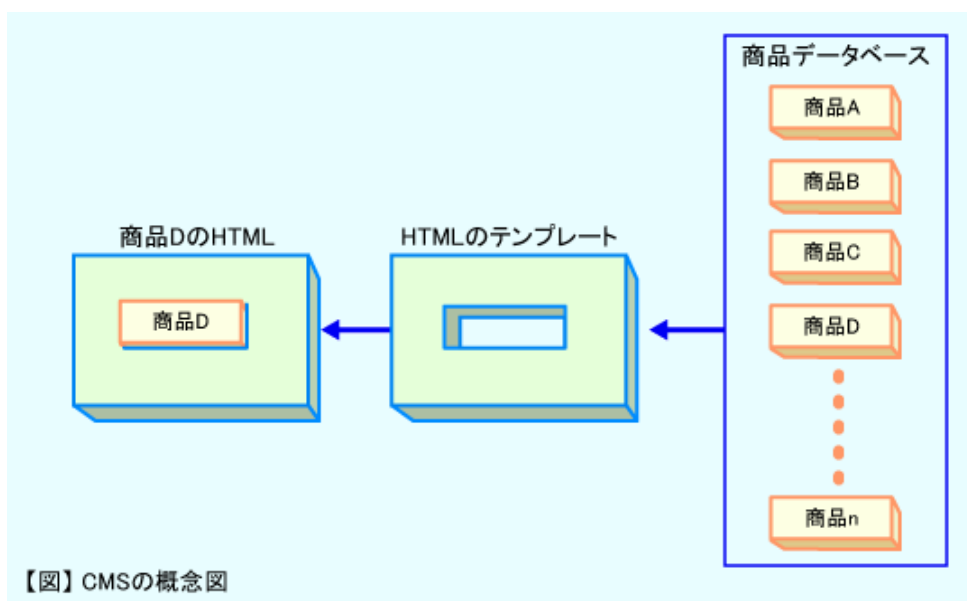
■ネットショップに見る Web サーバとデータベース

前述したサーバやデータベースが Web サーバでどのように働いているか、ネットショップを例にその様子を概観してみましょう。

・ネットショップの売場は、その都度作られる

商店の売り場スペースに当たるのがネットショップの商品紹介ページです。商品が多くなればページ数も多くなります。個々のページにはリンクを張ったボタンがあったり、一覧表示、詳細表示、おすすめ商品表示などを別々に作っていたりするのは、更新、保守、運用が大変です。そこで考えられた仕組みが CMS (Contents Management System、コンテンツ管理システム) です。表示用の HTML ファイルを完全な形で Web サーバに用意しておくのではなく、ブラウザからの要求に応じて HTML を組み立てて送信する仕組みです。組み立て用の部品は、テキスト(文字)や画像、HTML や CSS (注 1)などをバラバラにした状態でデータベースに保存しておきます。商品紹介ページの要求があれば、専用の HTML テンプレートに、データベースから取得した商品データを組み合わせて「商品紹介ページ」を組み立ててブラウザに送信します。同じ商品データを別のテンプレートと組み合わせて「今月のおすすめ商品」、「商品一覧表」を組み立てることができます。テンプレートを変えるだけで、同じ商品データがさまざまなページで利用できるようになります。

次の図はネットショップの CMS の概念図です。データベースから情報を抽出し、その結果とあらかじめ用意した Web ページのテンプレートを組み合わせて HTML を生成しています。テンプレートはそれぞれ必要なページ用に用意しておけば、商品をさまざまな形で見せることができます。

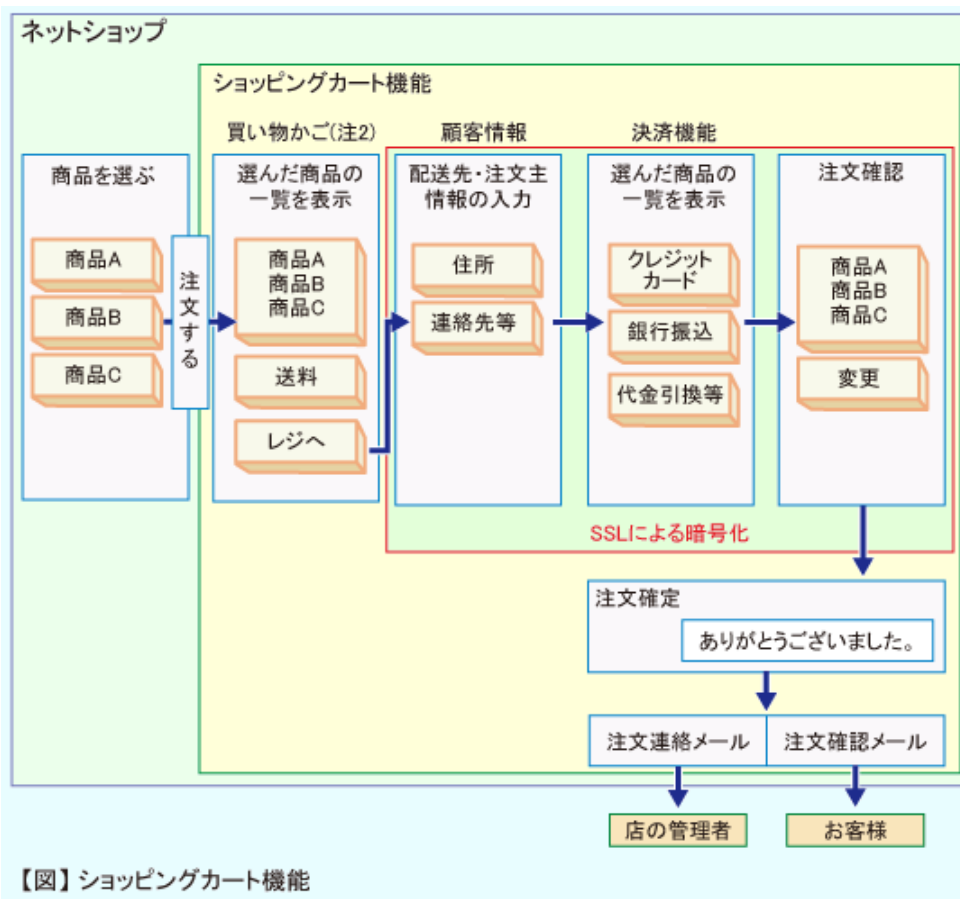


CMS は、ブラウザの要求に応じてその都度 HTML を作成します。このようにしてその都度作られる HTML を動的コンテンツまたは動的 HTML と呼んでいます。これに対して、Web サーバにあらかじめ作成して保存されているコンテンツを静的コンテンツあるいは静的 HTML と呼びます。今日の Web2.0 の世界では、静的なコンテンツでは表現力や可能なサービスに限界があります。動的なコンテンツは今や必須となっており、CMS はこうした Web サイトを支える中心的な仕組みです。

注1 CSS(Cascading Style Sheets) : Web ページのフォント、サイズ、画像やレイアウトなどのデザイン表現を記述したファイルで、HTML とセットで使用される。

・ネットショップの重要な仕組み、買い物かご

商店のキャッシャーのはたらきに相当するのが、ネットショップのショッピングカート（買い物かご）です。注文の確認から代金の支払いや発送方法の指定など一連の購入手続きのすべてを行うネットショップ用のプログラムです。次の図のような機能を持っています。



注2：機能としてのショッピングカートと区別するために、図中では買い物かごと表記しています。

ネットショップでの買い物は、サイト内の複数のページを行き来しながら、商品を選び、ショッピングカートに入れていきます。ショッピングカートが機能するためには、個々のお客様のブ

ブラウザを識別し、複数のページにまたがって継続的に情報を記憶しなくてはなりません。これを実現するのが前々回説明した Cookie です。ブラウザの識別や選んだ商品の履歴、表示したページの情報などがブラウザの Cookie に保存され、この情報をもとに、ショッピングカートによる買い物が可能になります。当然ながら、ブラウザの Cookie の設定はオンになっていないと、ショッピングカートは正しく機能しません。

ショッピングカートプログラムは、ブラウザで商品を選択するたびにデータベースから商品データを読み込み、アプリケーションサーバ内に記憶します。ショッピングカートの中味を確認するボタンをクリックすると、記憶している商品データとテンプレートと組み合わせて選択した商品の一覧表を作り、ブラウザに送信します。お客様名や住所などの顧客情報もデータベースに格納されていることがあり、アプリケーションサーバは入力内容をデータベースに問い合わせたり、データベースに記録したりしています。

「一覧表示」から最後の「注文確定」画面まですべて、データベースからの情報とテンプレートを、ショッピングカートプログラムが組み合わせて HTML を生成しています。CMS の 1 種と言えるでしょう。今日では、Web サイトとデータベースはきってもきれない間柄になっています。ショッピングカートなどのアプリケーションが充実してきて、データベースを意識することなく利用できるようになってきました。Web サイトの中でも特にネットショップは、商品紹介から顧客情報、在庫管理、発送情報など取り扱う情報量が多く、取扱商品数にかかわらず手動での管理はほとんど不可能な状態になっています。

■ これからの Web サイトとデータベース

今日の Web サイトには、情報の新鮮さ、量の充実、使いやすさ、対話性、楽しさ、快適な速度、安定した動作等々、多様な機能や性能が求められています。閲覧だけの静的なコンテンツの時代から、動的なコンテンツで Web を使う時代に移りました。その変化を先取りし、あるいはその波に乗って成長してきたのがネットショップです。1995 年 7 月に操業の Amazon.com は四半期で 28 億 9000 万ドル (2007 年第 2 四半期) を売り上げるまでに成長しました。膨大な取扱商品と顧客数、さまざまなおすすめ商品機能、ワンクリックショッピングなどのネットショップ機能はデータベースと CMS による動的コンテンツの生成技術なしでは成り立ちません。

Web サーバも進化しています。Web サーバ自体が、すでに単なるサーバソフトウェアではなくなりつつあります。Web サイトは単に静的なページを表示するだけではありません。ビジネス向けであれ消費者向けであれ、Web サイト自体がアプリケーションへと進化を遂げています。情報の検索や商品の購買は当然のもので、ページのカスタマイズや以前の情報を記憶しているサイトも珍しくなくなりつつあります。ファイルを送信するための環境から、アプリケーションプラットフォームへと進化を遂げつつあるのが今の Web サーバです。

また、現在データベースの大半がリレーショナルデータベースですが、今後、増えていくと言われているのが XML データベースです。リレーショナル型はデータがなくても空のセルを用意する必要があり、その意味では Excel の表に似ています。これに対して、テキストベースの XML は、空セルのような空白のデータを必要とせず、タグによって不定形の階層構造を表現でき、自

由度と効率の良さから利用が拡大しつつあります。XML は、今日のブログなどの隆盛を下から支えている技術であり、B2B では電子情報交換の標準となりつつあります。データベースの XML 化は、ネットショップをはじめ企業の Web サイトを大きく変えていく可能性があります。

今日、動的コンテンツを使用した Web サイトをごく普通に見かけるようになりました。これらのサイトを動かしているのが、今回お話しした三層構造システムです。今後、マッシュアップのような他システムとの高度な連携や複雑な Web サービスが実現されるにつれ、こうした Web サイトのしくみはますます磨きがかかってくると考えられます。

■おさらい

動的コンテンツとはブラウザからの要求に応じて Web サーバ内で組み立てられるコンテンツ
三層構造システムは、Web サーバ、アプリケーションサーバ、データベースサーバで成り立つ
データベースは、情報を統合・整理・蓄積し、便利に使う仕組み
データベースの三要素は、データベース本体、DBMS、SQL

・参考リンク

データベースと連携した Web ページを作成するための基礎知識

<http://park18.wakwak.com/~little-box/Dreamweaver/master043.htm>

ショッピングカートを設置しよう

<http://allabout.co.jp/career/hp4smallbiz/closeup/CU20070810A/>

データベースのイロハ：そもそもデータベースとは？ - builder by ZDNet Japan

<http://builder.japan.zdnet.com/sp/07database/story/0,3800082819,20355123,00.htm>

特集:基礎から理解するデータベースのしくみ：ITpro

<http://itpro.nikkeibp.co.jp/article/COLUMN/20060127/228070/?ST=middleware>