

## Möglichkeiten und Grenzen einiger frei verfügbarer Softwaresysteme für den Datenbankzugriff über WWW

P. GIEBLER, J. SPILKE, D. CONRAD, R. RAEKE, P. RIEBENSTAHL, HALLE

### 1. Einführung

Die enorme Entwicklung des WorldWideWeb (WWW) in den letzten Jahren erzeugt einen wachsenden Bedarf, Informationen nicht nur zu beziehen, sondern auch anzubieten.

Durch die allgemeine Verfügbarkeit der zugehörigen Software und deren einfache Nutzungsmöglichkeiten sowie immer leistungsfähigere und preisgünstigere Hardware, auch und vor allem im PC-Bereich, ist es prinzipiell für jeden möglich, Informationsanbieter im WWW zu werden.

Dabei nimmt neben dem WWW-Angebot im Internet auch die Nutzung der Internet-technologien für unternehmens- oder organisationsinterne Zwecke, dem sogenannten Intranet, an Bedeutung zu.

Gerade die freie Verfügbarkeit der benötigten Software kann für Institutionen, Einrichtungen und Betriebe von Interesse sein, die u.a. durch folgende Eigenschaften gekennzeichnet sind:

1. Sie besitzen keine eigene IT-Abteilung, die Aufbau, Betrieb, Wartung der WWW-Server professionellen Zuschnitts (Hard- und Software) permanent übernimmt.
2. Es erfolgt keine Delegation von Betrieb und Wartung des WWW-Servers an Dritte, weil die Kosten-Nutzen-Relation als zu ungünstig angesehen wird.
3. Es erfolgt ein Probetrieb eines WWW-Servers, um zunächst Erfahrungen zu sammeln.
4. Es ist absehbar, daß das Angebot von WWW-Inhalten oder die Nachfrage danach weder die Einrichtung einer eigens damit betrauten IT-Abteilung und die Beschaffung zugehöriger Hard- und Software noch die Delegation an Dritte erfordert.
5. Es ist ausschließlich PC-Technik verfügbar.

Beim Bereitstellen von Informationen im Internet bzw. Intranet ist im wesentlichen zu unterscheiden zwischen statischen Inhalten, die wenig und in großen Zeiträumen zu ändern und einfach zu warten sind und dynamischen Inhalten, also häufig geänderten, an den aktuellen Stand der internen Datenversorgung der Einrichtung angepaßte Inhalte. Diese können insbesondere auch WWW-Inhalte sein, die sich durch Abfrage von Datenbanken ergeben und die auch die Form der angebotenen WWW-Seiten beeinflussen.

Gerade die Bereitstellung dynamischer Inhalte und die Entwicklung dazu nötiger Technologien sind eine wesentliche Entwicklungslinie innerhalb der WWW-Technologie der letzten Jahre.

Dazu existieren mittlerweile eine Vielzahl professioneller Entwicklungssysteme, z.T. auch angebunden an leistungsfähige Datenbanksysteme. Aber auch unterhalb der Ebene, die durch professionelle Produkte des Zugriffs durch WWW-Server auf Datenbanken abgedeckt wird, existiert eine Vielzahl von Ansätzen und praktikablen Lösungen, die wenig finanziellen Aufwand benötigen oder kostenlos verfügbar sind. Die Leistungsfähigkeit ist im Vergleich zu kommerziellen Produkten aber meist entsprechend geringer.

Im Hinblick auf eine Nutzung in oben charakterisierten Einrichtungen ist somit zu untersuchen, welche Ergebnisse mit frei verfügbaren Softwaresystemen zur Datenbankanbindung an WWW-Server zu erzielen sind. Bei diesem Vergleich sollen technischer und personeller Aufwand im Vordergrund stehen.

### 2. Technische Möglichkeiten und Lösungen

WWW-Server sind im allgemeinen nicht in der Lage, andere als speziell formatierte Dateien, die HTML-Dokumente, verfügbar zu machen. Ein Zugriff auf Datenbanken ist nicht direkt möglich. Meist sind derartige Daten durch Zugriffssysteme gesichert und nur über von einem Datenbanksystem kontrollierte Transaktionen verfügbar. Zur Verbindung zwischen WWW-Server und Datenbank sind Gateways nötig, die diese Transaktionen organisieren. Bedingt durch die technische Entwicklung gibt es sehr unterschiedliche Lösungen derartiger Gateway-Technologien.

Gemeinsam ist ihnen lediglich, daß sie nach Anforderung an einen WWW-Server die Verbindung zwischen diesem und einem Datenbestand bzw. einer Datenbank herstellen, die dort abgefragten Daten formatieren und dem WWW-Server zur Lieferung an den WWW-Client übergeben.

Zwei wesentliche Grundprinzipien lassen sich jedoch unterscheiden:

- **CGI (Common Gateway Interface)**  
ist die Standard-Schnittstelle der WWW-Server, um zur Abfragezeit beliebige Programme (Prozesse) zu starten und z.B. HTML-Dateien zu erzeugen. Dabei ist es (in Unix) völlig gleichgültig, ob die Programme aus C oder einer anderen Programmiersprache kompiliert, aus einer Scriptsprache wie Perl oder als Unix-Shell-Scripts vorliegen. Mittlerweile besitzen auch die Windows-basierenden WWW-Server eine CGI-Schnittstelle, wobei unter Windows (Win95/98, WinNT) andere Lösungen zu bevorzugen sind, falls nicht auf eine betriebssystemunabhängige Lösung Wert gelegt wird.
- **Server-Erweiterungen durch API (Application Programming Interface)** sind (proprietäre) Low-Level-Schnittstellen, die die Funktionalität eines Servers (unter MS Windows) anwendungs- oder aufgabenspezifisch erweitern. Ein Vorteil derartiger Servererweiterungen ist, daß sie meist im selben Prozeß-/Adreßraum ablaufen wie der WWW-Server selbst, also nicht wie bei CGI für jede Abfrage ein separater Prozeß gestartet werden muß. Nachteilig ist die Plattformabhängigkeit durch die starke Bindung an den jeweiligen Hersteller der API und die damit verbundene geringe Portierbarkeit.

Der Zugang zu den Datenbeständen wird unter MS-Windows meist durch **ODBC (Open Data Base Connectivity)** hergestellt, welches eine API-Schicht über (beliebige) Datenquellen legt. Zu den meisten (PC-)Datenbanksystemen existieren ODBC-Treiber. Diese stellen die Verbindung zwischen Anwendungsprogrammen (z.B. WWW-Servern) und der Datenquelle her, die von Seiten des WWW-Servers dadurch weitgehend einheitlich erscheint. Dadurch wird es möglich, ein (z.B. zunächst versuchsweise genutztes) Desktop-System wie MS-Access bei steigender Beanspruchung durch ein leistungsfähigeres Serversystem (z.B. MS-SQL-Server) zu ersetzen, ohne daß die WWW-seitige Anbindung wesentlich geändert werden muß. Die Standardabfragesprache für Datenbanken über ODBC ist SQL.

Einige frei verfügbare Server-Erweiterungen unter MS-Windows (mit unterschiedlicher Komplexität) sind:

- **hotSQL** - ein sehr einfaches, unter Windows (Win95, WinNT) über CGI einsetzbares WWW-DB-Gateway
- **MS dbWeb** - ISAPI-dll zu MS IIS, zur Erstellung von Web-Seiten zur Abfrage von Datenbanken und zur Realisierung dieser Abfragen
- **IDC/HTX (Internet Database Connector/ HTML-Extension)** - eine bereits im MS-IIS 1.0 enthaltene ISAPI-DLL mit Namen httpodbc.dll, die den Zugriff über ODBC auf Datenbanken ermöglicht.

- **ASP** (Active Server Pages) sind eine Erweiterung des MS IIS 2.0. Sie stellen ein Anwendungs-Framework dar, welches neben einer vollwertigen Programmiersprache auch anwendungsspezifische Objektklassenbibliotheken enthält. Der Leistungsumfang ist der eines professionellen Systems. Dabei sind Kenntnisse der strukturierten bzw. objektorientierten Programmierung nötig.

Für eine Datenbankanbindung über WWW nicht direkt notwendig, aber zu Formatierungs- und Gestaltungszwecken meist sehr sinnvoll einsetzbar sind **Scripts** (Javascript, VB-Script, Perl). Die Nutzung von Scriptsprachen ermöglicht z.B. eine Formatierung der Ein- und Ausgabe bzw. eine über die in HTML enthaltenen Sprachmittel hinausgehende Formulgestaltung. Die Nutzung von ASP setzt die Verwendung von Scripts voraus.

### 3. Tests und Bewertungen

Die zu realisierenden dynamischen WWW-Angebote waren beschränkt auf die folgenden „einfachen“ Aufgaben:

- Anzeige von Datenbankinhalten nach einem oder mehreren Auswahlkriterien, (z.B. Abfrage eines Veranstaltungskalenders oder Angebotslisten, die nach vorgegebenen Kriterien eingeschränkt werden)
- Erfassen von Datensätzen in einer Datenbank (Einträge in Formulare).

Dabei waren verschiedene Aspekte zu betrachten:

- Leistungsumfang und damit Eignung für die jeweilige Aufgabe
- Reaktionsverhalten, Antwortzeiten bei Abfragen
- Einarbeitungsaufwand
- Aufwand für Betrieb, Änderungen und Wartung.

#### Untersuchte Konfigurationen

|  | Konfiguration I        | Konfiguration II     | Konfiguration III                | Konfiguration IV                 |
|--|------------------------|----------------------|----------------------------------|----------------------------------|
| <b>Computer</b>  | P75/64MB               | PII200/128MB         | PII200/128MB                     | PII200/128MB                     |
| Hardware und Betriebssystem  | MS Windows95           | MS Windows NT Server | MS Windows NT Server             | MS Windows NT Server             |
| <b>WWW-Server</b>  | MS Personal Web Server | MS IIS               | MS IIS                           | MS IIS                           |
| <b>Datenbank-anbindung</b> und Generierung dynamischer HTML-Seiten | HotSQL                 | dbWeb                | IDC/HTX (integriert in MS IIS)   | ASP (integriert in MS IIS)       |
| <b>Datenbanksystem</b>   | MS Access              | MS Access            | MS Access (a), MS SQL-Server (b) | MS Access (a), MS SQL-Server (b) |

Zum Vergleich wurde ein professionelles System (SQL-Surfer von Netaway, ergänzt durch Java-Script) anstelle von IDC/HTX in Konfiguration III verwendet.

Die Netzverbindung erfolgte über LAN (Ethernet 10 Mbit/s) als auch per ISDN/Modem.

Nicht enthalten in den betrachteten Lösungen waren Zugriffsprotokolle, Sicherheitsvorkehrungen und die Möglichkeiten der Fernadministration.

#### Leistungsumfang und Eignung

- **hotSQL** (Konfig. I) vermittelt lediglich einen Eindruck über die prinzipielle Wirkungsweise einer Datenbankanbindung über WWW. Für weitergehende Zwecke ist der Funktionsumfang zu gering.
- **dbWeb** ist einsetzbar für einfache, sich wenig ändernde Aufgaben; variable, von einem vorgegebenen Schema abweichende Abfragen sind schwer oder garnicht zu realisieren.
- **IDC/HTX** ist komfortabler, wobei aber schon die Strukturierung der Ergebnispräsentation ohne den Einsatz von Scripts (z.B. Java oder VB) nicht möglich ist.
- **ASP** als System auch für den professionellen Einsatz besitzt einen Funktionsumfang, der über die Anforderungen im Sinne oben beschriebener Einrichtungen hinausgeht.

#### Reaktionsverhalten, Antwortzeiten bei Abfragen

Von Interesse waren sowohl die Reaktionszeiten bei einzelnen Abfragen als auch bei mehrfachem gleichzeitigen Zugriff. Um eine Vergleichbarkeit zu sichern, wurde eine Test-Datenbank (in MS Access) generiert, die ausreichend groß und komplex war, um Abfragen mit unterschiedlichen Antwortzeiten zu ermöglichen.

Bei der Verwendung von Access existiert eine Abhängigkeit der Zugriffszeit von der Komplexität der Abfrage (sowohl der Zugriff über IDC/HTX zur MS-Access-Datenbank als auch bei der vergleichsweise durchgeführten Abfrage in MS Access selbst)

Die gleiche, nach MS SQL-Server 6.5 importierte Datenbank, ergibt keine merkbaren Unterschiede in der Zugriffszeit.

Der Zugriff über ISDN (oder Modem) zeigte sich eher abhängig von der Qualität der Verbindung als von der Komplexität der Abfrage, denn bei höherer Netzbeanspruchung oder langsamer Modem-Verbindung ist kein Zusammenhang zwischen Antwortzeit und Komplexität der SQL-Abfrage erkennbar.

Da der ODBC-Treiber von Access lediglich jeweils eine Abfrage behandeln kann, muß bei Mehrfachabfrage („fast“ gleichzeitig eintreffende Abfragen beim WWW-Server) die zuletzt eingetroffene Abfrage z.T. so lange warten bis der Anwender wegen time-out mitgeteilt bekommt, daß der gefragte Server nicht antwortet. Eine solche Konstellation ist offenbar nur dann verwendbar, wenn abzuschätzen ist, daß wenige Zugriffe gleichzeitig erfolgen werden. Derartige Effekte treten bei einem Client-Server-Datenbanksystem wie MS SQL-Server in Konfiguration IIIb nicht auf (auch nicht beim zum Vergleich getesteten SQL-Surfer).

#### Aufwand für Einarbeitung oder Schulung

Bis auf ASP ist der Umgang mit den Systemen relativ leicht erlernbar.

Vorausgesetzt wird dabei die Fähigkeit des Umgangs mit HTML-Editoren, Web-Site-Management-Tools, den zu nutzenden Datenbanksystemen, dem Betriebssystem und der Hardware.

Falls Scriptsprachen genutzt werden sollen, sind Erfahrungen darin günstig bzw. in einer Programmiersprache wie C++, Pascal oder Fortran, wodurch die Einarbeitung in die Scriptsprachen erleichtert wird.

#### Aufwand für Systempflege und Wartung

Ohne Werkzeuge, die eine Systemintegrität sichern helfen (dbWeb, ASP mit MS Frontpage), ist der Aufwand für Systempflege und Wartung ab einer gewissen Komplexität unzumutbar hoch.