Serverbasierte Programmiersprachen im WWW

Bastian Friedrich

Seminar

„Anwendungsorientierte Programmiersprachen“

WS 2000/2001

Freiburg, Februar 2001

Übersicht

- Problembeschreibung
- Lösungswege
- Sprachen
- Fazit
- Diskussion

Programmiersprache für das WWW (1)

- Einfachheit
- Stabilität, Sicherheit
- Datenbanken
- Flexibilität

Programmiersprache für das WWW (2)

- Reflexion der Belange des WWW
  - XML/HTML
  - Forms
- Integration in den Webserver
- Entwicklungstools

Lösungswege

- CGI
- FastCGI (FCGI)
- Anwendungsspezifische Sprachen
  - Servlets/JSP
  - PHP
  - ASP/VBS (Active Server Pages)
  - <bigwig>
  - Andere

Allgemeine Schnittstellen: CGI

- Common Gateway Interface
  - Nimmt der Programmiersprache das „Drumherum“ ab: HTTP-Header, ...
  - Beliebige Programmiersprache
  - Datenaustausch über Environment und Standard-Ein-/Ausgabe
  - fork() für jedes CGI
    - Mäßige Performance
  - Immernoch Standard für die meisten Projekte

Allgemeine Schnittstellen: FastCGI

- Fast Common Gateway Interface
  - Basis: CGI
  - Unterschiede
    - Persistente Prozesse
    - Persistente Bindungen (Datenbanken, Dateien)
  - In der Tat: Fast!
  - Programmierung komplexer und sensibler als CGI
  - Nicht mehr Teil von Apache

```
#!/usr/bin/perl

use FCGI;
$count = 0;

while (FCGI::accept() >= 0) {
  print('Content-type: text/html\r\n\r\n',
        'FastCGI sagt zum ',++$count,
        '-ten mal Hallo!\n');
}
```

Sprachen: Perl/ePerl (1)

- Perl 1.0 in 1987 von Larry Wall
- Schnelle Entwicklung bis 1994: Perl 5
- Perl 6
  - Komplette Neuimplementierung
    - „Gesundschrumpfung“
    - Integration moderner Konzepte (Corba, ...)
    - Typisierung?
  - Implementierungsphase - oder auch nicht

Sprachen: Perl/ePerl (2)

- Unterstützung des WWW
  - spezielle Module (CGI.pm)
  - Einfache Stringbehandlung
- Typsystem
  - Untypisiert
    - Für kleine Projekte sehr angenehm
    - Bei größeren Projekten: Zunehmend schwerer zu handhaben
  - Objektbasiert

Sprachen: Perl/ePerl (3)

- Viele frei verfügbarer Bibliotheken (CPAN)
  - Code-recycling
  - „Plug'n'Play"
- Stringbehandlung
  - dynamische Speicher-Allozierung
    - Verhinderung von Buffer-Overflows
  - Regular Expressions
- Einfache Ausgabe und Auswertung von Forms über CGI.pm

Sprachen: Perl/ePerl (4)

- Precompiling
- ePerl: Einbettung von Perl-Kommandos direkt in HTML-Text
  - Schnelle Entwicklung
  - Keine Trennung von „View“ und Algorithmen

Sprachen: Perl/ePerl (5)

- Sicherheitsprobleme:
  - SQL-Quoting
  - Voller Zugriff auf das Filesystem - evtl. „elevated privileges“
  - format string bugs
  - race conditions
  - „suidperl“ - heute Problemlos?

Sprachen: Perl/ePerl (6)

- Sicherheitsfeatures
  - String-Behandlung
  - Verifizierte Module
  - Große Entwicklergemeinde - gut getestet
  - Taint-Checking
    - offensichtliche sowie subtile Sicherheitschecks

Sprachen: Perl/ePerl (7)

- Precompiler verbessern die Geschwindigkeit kaum
- mod_perl, FastCGI gegen forking
- Profiler im Paket enthalten

```
#!/usr/bin/perl

use CGI qw/:standard/;
print header,
      start_html('A Simple Example'),
      h1('A Simple Example'),
      start_form,
      "What's your name? ",textfield('name'),p,
      "What's the combination?", p,
      checkbox_group(-name=>'words',
                     -values=>['eenie','meenie','minie','moe'],
                     -defaults=>['eenie','minie']), p,
      "What's your favorite color? ",
      popup_menu(-name=>'color',
                 -values=>['red','green','blue','chartreuse']),p,
      submit,
      end_form,
      hr;

 if (param()) {
     print "Your name is ",em(param('name')),p,
           "The keywords are: ",em(join(", ",param('words'))),p,
           "Your favorite color is ",em(param('color')),
           hr;
 }

 print end_html;
 print "\r\n";
```

Sprachen: C/C++ (1)

- Entwicklung von C: Brian Kernighan/Dennis Richie Anfang der 1970er
  - Standardisierung 1983
- Entwicklung von C++: Bjarne Stroustrup, Anfang der 1980er
  - Standardisierung 1998

Sprachen: C/C++ (2)

- Keine inhärente Web-Fähigkeiten
  - Freie Bibliotheken verfügbar
  - Schnell
  - Viele „andere“ Bibliotheken
  - Große Entwicklergemeinde

Sprachen: C/C++ (3)

- Typsystem
  - Sprache benötigt häufige casts
  - Typsicherheit gefährdet
  - Besserung in C++
- Sicherheitsprobleme
  - Systemnähe und mangelnde Kapselung
  - Praktisch alles, was an Problemen denkbar ist, ist über C auszunutzen
- Performance: Gut, aber CGI...

Sprachen: PHP3/4 (1)

- Personal Home Page Tools 1995
- PHP3: 1997; PHP4: 2000
- Findet schnell Verbreitung:
  - Verfügbar für
    - Apache
    - IIS
    - Netscape/AOL
    - Roxen
    - ...

Sprachen: PHP3/4 (2)

- Als Web-Sprache entwickelt
  - „Embedded“-Sprache: HTML mit Programm vermischt
  - Direktes Handling von POST/GET-Variablen und Cookies
  - Einfaches Setzen von Header-Informationen
  - Integration von Datenbankfähigkeiten

Sprachen: PHP3/4 (3)

- Probleme in der Web-Integration:
  - Kaum XML-Unterstützung
  - Keine integrierte Unterstützung von HTML-Generierung
    - Formular-Generierung?
    - HTML-Korrektur?
    - ...

Sprachen: PHP3/4 (4)

- Sprache stark an Perl angelehnt
- Untypisiert
  - Objektbasiert
  - Wie Perl
    - für kleine Projekte angenehm
    - Bei zunehmender Projektgröße kaum noch zu handhaben

Sprachen: PHP3/4 (5)

- Features
  - Anbindung an beliebige Datenbanken über ODBC
  - Anbindung an viele Datenbanken über eigene Treiber
  - Server-Integration über Module
  - Regular Expressions
  - Integrierte Grafik-Generierung (gdlib)
  - Integrierte Kryptographie (mcrypt)
  - ...

Sprachen: PHP3/4 (6)

- In der Vergangenheit mehrfach Sicherheitsprobleme
- Safe-Mode
  - Kein Zugriff auf Daten außerhalb des Doc-Root
- AddSlashes
- HtmlEntities
- Trotz der Features: Aufpassen!

Sprachen: PHP3/4 (7)

- PHP4 mit neuer Engine
  - Performance-Steigerung um bis zu 200%
- Integration in den Server: kein Forking
- Trotz Interpretation: Zügig
- Bis vor kurzem: Keine Connection-Pools für Files und DB
- PHP4: Caching, ... - PHP wird erwachsen (?)

```
<html>
  <head>
    <title>Beispiel 2: PHP</title>
  </head>
  <body bgcolor="#FFFFFF">
    <h1>Beispiel 2: PHP</h1>
    <table>
    <script language="php">
      mysql_connect("localhost","nobody");
      mysql_select_db("vortrag");

      $res = mysql_query("SELECT * FROM vortrag ORDER BY
vorname");
      for ($i = 0; $i < mysql_num_rows($res); $i++) {
        $arr = mysql_fetch_array($res);

print("<tr><td>$arr[id]</td><td>$arr[vorname]</td><td>$ar
r[name]</td></tr>");
      }
    </script>
    </table>
  </body>
</html>
```

Sprachen: Java, Servlets, JSP (1)

- Produktfamilie auf der Java-Hype-Welle
- Servlets 2.1: 1998
- JSP 1.0: Ende 1999
- JSP 1.1: 3 Monate später

Sprachen: Java, Servlets, JSP (2)

- Erfordern direkte Unterstützung des Webservers
  - Referenzimplementation in Apache nach wie vor problematisch
  - JSP: Embedded-Sprache
  - Abfrage von POST/GET-Daten und Cookies mit dem üblichen Java-Overhead
  - Angeblich: XML-Konform
  - Datenbank-Zugriff über JDBC

Sprachen: Java, Servlets, JSP (3)

- Recht aufwändiges Setup
- Volle Nutzung der Java-Ressourcen
  - Ausreichend Bibliotheken vorhanden
- Keine eigenständige HTML-Generierung
- Keine Unterstützung von Client-Side-Scripting

Sprachen: Java, Servlets, JSP (4)

- Starke Typisierung
- Konsequent objektorientiert
  - Komponenten Bestandteil des Konzepts
    - Schlagwort „Enterprise Java Beans“

Sprachen: Java, Servlets, JSP (5)

- Sicherheitsfeatures
  - Java Sandbox
  - Java Security
    - ACLs
    - Zertifikate
  - Keine Pointer-Arithmetik
    - Keine Buffer-Überläufe
  - Keine Format-String-Exploits
  - Keine Spezielle „Web-Sicherheit“

Sprachen: Java, Servlets, JSP (6)

- Freie Implementierungen nach wie vor mangelhaft
- Äußerst sensible Konfiguration
- Kein „public_html“

Sprachen: Java, Servlets, JSP (7)

- Performance
  - Java-Typisch
  - Lange Kompilierungs-Zeit
  - Runtime-Environment persistent und zügig
  - Durch Persistenz: Connection Pooling möglich

Sprachen: ASP/VBS (1)

- (fast) Proprietär (MicroSoft/IIS)
  - Sekundär-Implementierungen kommerziell oder nur für Windows
- Benutzbar seit ca. 1997
- Moderne Schlagworte
  - Komponenten (COM/DCOM)
  - XML
  - Threading...

Sprachen: ASP/VBS (2)

- Embedded-Sprache (wie PHP und JSP)
  - Haupt-Sprache: Visual Basic Script
- Keine Generierung von HTML und Client-Side-Scripts
- Einfacher Zugriff auf Formular-Daten
- Zugriff auf Datenbanken über Komponentenmodell

Sprachen: ASP/VBS (3)

- VBS: Untypisiert
- Stärken
  - Zugriff auf externe Komponenten
  - Gute (Idee der) Entwicklungsumgebung

Sprachen: ASP/VBS (4)

- Sicherheit?
  - MicroSoft!
  - Wiederholte Sicherheitslöcher
    - Zugriff auf Script-Sourcen
    - Bis vor kurzem: Zu enge Integration der Runtime-Umgebung
- Performance
  - Integration in den Server
  - Zunehmend besser

Sprachen: <bigwig> (1)

- Dänisches Universitätsprojekt seit 1998
  - „Geistiger Nachfolger“ von MAWL
  - Zielgruppe: Normale Programmierer
  - Kompilierte Sprache
  - Sitzungsorientierte Programmierung
    - Pseudo-Persistente Daten

Sprachen: <bigwig> (2)

- Ergebnis
  - HTML
  - CGI-Scripts
  - JavaScript (kleinster gemeinsamer Nenner von Internet Explorer/Mozilla)
  - HTTP Authentication
  - Java Applets
  - Offen für neue Technologien

Sprachen: <bigwig> (3)

- Direkte Integration aller aktuellen Web-Technologien
- Keine inhärente Unterstützung von „ungeplanten“ Technologien (Flash, ...)
- Heute: Kein XML
- Datenbank-Anbindung: Eigene Typen für Relationen und Tupel

Sprachen: <bigwig> (4)

- Typisierung
  - starke Typisierung
    - Grundgedanke: Anlehnung an Java, C, ...
- Besonderheiten
  - Pseudo-Persistente Sitzungen („Sessions“)
  - Generierung von Client-Side-Scripts
  - Konsequente Trennung von Code, Daten und „Views“

Sprachen: <bigwig> (5)

- „Security - We have not fully analyzed this enormous aspect, ...“
- Trennung vom Web-Server
  - Layer-Technologie
- Keine Pointer-Arithmetik
- Kleine Entwicklergemeinde

Sprachen: <bigwig> (6)

- Performance-Probleme von CGIs
- Nativer Code
- Projekt „<bigwulf>“
  - <bigwig> und Beowulf Clustering

Sprachen: Andere (1)

- WeSQL
  - Embedded SQL in HTML-Seiten
  - Von Apache durch Perl-Script geparsed
- Roxen
  - Eigene Programmier-Engine
- Haskell, Python, ...
  - Selten anzutreffen
  - Auch hier: Bibliotheken vorhanden!

Sprachen: Andere (2)

- CMS
  - Content Management Systems
  - Keine Programmiersprachen in diesem Sinne
  - Darstellung von Informationen, die von der Webseite gelöst gespeichert sind

```
<!-- LIST A SELECT name,id_somethingelse FROM somethings WHERE id='PR_ID|[0]' -->
<table width="100%" border=1 BGCOLOR="#F2F1FE">
<tr valign="middle"><td valign="middle"><center><font color="#000099"
size="+2">Something Information for "A_NAME".</font></center></td></tr></table><br>
<br>
<a href="/index">Home</a><br><br>
<table><tr><td>
<form method=post action="/modify">
<input type=hidden name=id value="PR_ID">
<input type=hidden name=table value="somethings">
<input type=hidden name=redirdest value="/">
Name:
<input name=name value="A_NAME">
Related to:
<select name=id_somethingelse>
<!-- LIST B SELECT id,name FROM somethings WHERE id='A_ID_SOMETHINGELSE' ORDER BY name
-->
<option value="B_ID">B_NAME</option>
<!-- /LIST B   -->
<!-- LIST B SELECT id,name FROM somethings ORDER BY name -->
<option value="B_ID">B_NAME</option>
<!-- /LIST B   -->
</select>
<input type=submit value="Update" name="button">
</form>
</td><td>
<form method=post action="/del">
<input type=hidden name=id value="PR_ID">
<input type=hidden name=table value="somethings">
<input type=hidden name=redirdest value="/">
<input type=submit value="Delete" name="button">
</form>
</td></tr>
</table>
<!-- /LIST A <H1>There is no such Something with id 'PR_ID|[0]'</H1> -->
```

Mein Fazit (1)

- Mittlerweile durchaus Auswahl vorhanden
- Perl: Stellung gerechtfertigt
- C/C++: Zu sensibel; keine Eignung für kleine Projekte
- PHP3/4: Hype - schnelle Erfolge bei kleinen Projekten
- JSP: Für die meisten User Vaporware

Mein Fazit (2)

- ASP: moderne Schlagworte machen keine moderne Sprache!
- <bigwig>: Interessante Ideen; zu wenig Entwickler und Informationen
- WeSQL: nicht mehr als Kleinigkeiten
- Roxen: Proprietär
- Haskell, Python, ...: Nur in Spezial-Fällen!

Diskussion

- Fragen?
- Meinungen?
- Erfahrungen?