

Web Application Software

Jeff Offutt

<http://www.cs.gmu.edu/~offutt/>

SWE 432

Design and Implementation of
Software for the Web

“Where is the wisdom we lost in knowledge?
Where is the knowledge we lost in information?”
– T.S. Eliot

Motivation – Overview

- Modern web applications are:
 - Distributed (world-wide)
 - Heterogeneous (hardware and software)
 - Highly user interactive
 - Built on new technology
- The software is:
 - Very loosely coupled
 - Written in multiple languages
 - Often generated dynamically

Diverse: In terms of software, communication, and people

Motivation – Overview (2)

- Web site software has to be better than most shrink-wrap or contract software
- The combination of higher quality requirements and unique technologies make for a very *interesting* situation
(Academics think “interesting” means fun, managers think “interesting” is scary ...)

This talk discusses why and in what ways web software must be better

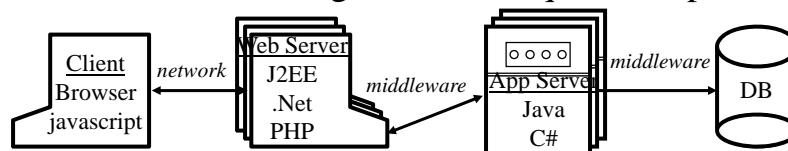
September 2, 2011

© Ofutt, 2001-2011

3

General Web Terminology

- Web Page : Data that fits in one browser screen
 - Static : HTML exists as a file on a computer
 - Dynamic : Created as needed
- Web Site : A collection of connected web pages
- Web Application : A program that is deployed on the web
 - UI is in HTML
 - User interacts through HTTP's request / response cycle



September 2, 2011

© Ofutt, 2001-2011

4

Software Deployment Methods

- Bundled : On your computer when you buy it
- Shrink-wrapped : Bought at a store on a CD
 - Downloaded from company’s website or OSS site
- Contract : Single customer
- Embedded : Installed on an electronic device
- Web application : On the web through a URL
 - Component-based
 - Concurrent / distributed
 - One copy on the server
 - Can be updated at any time (fast update cycle)
 - User interactive

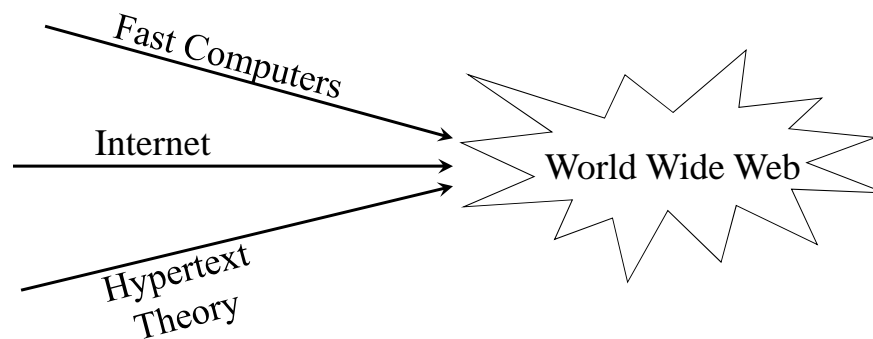
September 2, 2011

© Ofutt, 2001-2011

5

Formation of the WWW

The World Wide Web is a result of the
confluence of three developments



September 2, 2011

© Ofutt, 2001-2011

6

Historical Highlights

Hypertext, Internet, and the Web

- 1945: Vannevar Bush proposes hypertext
- 1965: Ted Nelson coins the term “Hypertext”
 - Douglas Engelbart invents mouse
- 1968: First hypertext systems (Engelbart’s “Augment/NLS”, Brown university’s HES)
- 1969: Arpanet
- 1970: Unix
- 1972: Tomlinson invents email
- 1978: First bulletin board system – Christensen and Sues
- 1982: < 100 sites on “pre-internet” (netnews)
- 1983: TCP / IP defines the internet
- 1986: NSFNET, SGML
- 1987: HyperCard (Apple), first hypertext workshop

September 2, 2011

© Ofutt, 2001-2011

7

Historical Highlights (cont.)

- 1989: WWW proposed by Berners-Lee
- 1990: HTML defined
- 1992: CERN (Switzerland) releases WWW
- 1993: First browser: NCSA Mosaic
- 1994: First widely used commercial browser: Netscape
- 1995: Lycos search engine, Java, DEC’s Alta Vista search engine
 - 15 million pages
- 1997: 31,000,000 pages, 476,000 servers, 14,000 news groups
- 2000: More than 100,000,000 hosts
 - Amount of back-end programming exceeds the front-end hypertext
- 2004: 3,307,998,701 pages (google), 500,000,000 hosts,
 - 1,200,000,000 users
 - Facebook launches
- 2006: Twitter launches
- 2010: Borders bankrupt, largest use of the internet is porn and games

September 2, 2011

© Ofutt, 2001-2011

8

Historical Highlights (cont.)

1994: WWW surpasses gopher

1994: Netscape

1995: Lycos search engine

Java

DEC's Alta Vista search engine

15 million pages

1996: US Telecommunications Bill

1997: 31,000,000 pages, 476,000 servers, 14,000 news groups

2000: More than 100,000,000 hosts

Amount of back-end programming exceeds the front-end
hypertext

2004: 3,307,998,701 pages (google), 500,000,000 hosts,
1,200,000,000 users

September 2, 2011

© Ofutt, 2001-2011

9

General Hypertext Terminology

- Node : Atomic amount of text – smallest unit of information known to system (pages)
- Link : Connects one node to another. Depicted by buttons or highlighted region.
 - Link Region (anchor) : Location in a node where link starts.
- Traverse (visit or view) : Moving from one node to another
 - Backtrack : Following links in reverse
- Landmark : Semantically important nodes that are linked to many nodes
- Hypermedia : Incorporates sound, video, pictures, or other non-textual information

September 2, 2011

© Ofutt, 2001-2011

10

The World Wide Web

The world wide web was invented by Tim Berners-Lee, an Englishman at CERN, the physics research lab in Switzerland

Tim Berners-Lee's vision of the world wide web was to make

- **all information available**
- **to all people**
- **all the time**

General Web Terminology

- Web Page : Data that fits in one browser screen
 - Static Web Page : A web page whose HTML exists as a file on a computer
 - Dynamic Web Page : A web page that is created as needed
- Web Site : A number of web pages that are *semantically* related and *physically* linked
- Web Site Software : Software that makes web sites dynamic.
 1. Dynamic Web Pages
 2. User Interaction

Modern Web Sites

- Web sites are now too complicated for individuals to manage
- They need to be engineered by teams of people with diverse talents:
 - Programming skills
 - Graphics design
 - Usability
 - Information layout and engineering
 - Data communications
 - Data base

We need web site engineering

September 2, 2011

© Ofutt, 2001-2011

13

Important Quality Attributes for Traditional Software

Traditional

1. Efficiency of process (time-to-market)
2. Efficiency of execution (performance)

⋮

50. Reliability

51. Safety

52. Maintainability

53. Security

⋮

September 2, 2011

© Ofutt, 2001-2011

14

Important Quality Attributes

1. Reliability
 2. Usability
 3. Security
-
4. Availability
 5. Scalability
 6. Maintainability
 7. Performance & Time to market

Customers have little “site loyalty” and will switch quickly, thus time to market is much less important than in other application areas.

(but still important!)

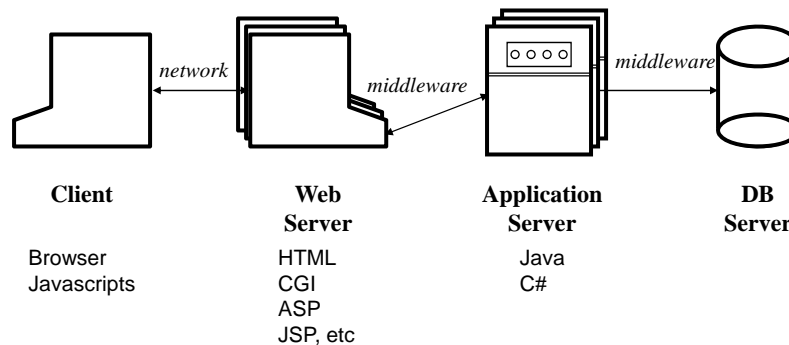
Based on an informal survey of around a dozen software development managers, 2000.

September 2, 2011

© Ofutt, 2001-2011

15

Fourth Generation Web Applications (software-centered)



Client-server ... 3-tier ... N-tier ...

September 2, 2011

© Ofutt, 2001-2011

16

Problems Can Occur Anywhere

- 1995: Web sites were 100% interface
- 1998: Web sites were about 90% interface
- 2001: Web applications are less than 50% interface
- 2005: Web applications about 25% interface
- 2011 : Web application development dominates the software industry

There is still a shortage of knowledgeable, skilled web programmers and software engineers

Summary Changes in Web Technologies

<u>1997</u>	<u>1998-1999</u>	<u>2000-2004</u>	<u>2004-2011</u>
<ul style="list-style-type: none">■ Static web pages■ “Soft brochures”■ Webmasters■ HTML, CGI, JavaScript	<ul style="list-style-type: none">■ Dynamic HTML■ Programs (poorly written)■ Confused webmasters■ ASP, CSS, ...	<ul style="list-style-type: none">■ Web applications■ ECommerce +■ Web services■ WebManager + programmers, DB, network, UI, graphics designers, ...■ Java (JSP, Servlets, beans), {HT,U,X}ML, Component-based■ .NET	<ul style="list-style-type: none">■ Web application frameworks■ Web 2.0■ AJAX■ Widespread enterprise applications

Summary Concerns of Software

Traditional

1. Efficiency of process
(time to market)
2. Efficiency of execution

○
○

50. Reliability
51. Safety
52. Maintainability
53. Security

○
○

Web Software

1. Reliability
2. Usability
3. Scalability
4. Security
5. Availability
6. Maintainability

○
○

September 2, 2011

© Ofutt, 2001-2011

19

Subject of 432

- 432 splits web software into two aspects:
 1. The interface aspect
 2. The software aspect
- The interface runs on the client side
 - Implementation of HTML is easy
 - Achieving usability is a key, and very difficult
- The software runs on the server side
 - A mixture of lots of technologies
 - The key to other quality criteria (reliability, etc.)

September 2, 2011

© Ofutt, 2001-2011

20