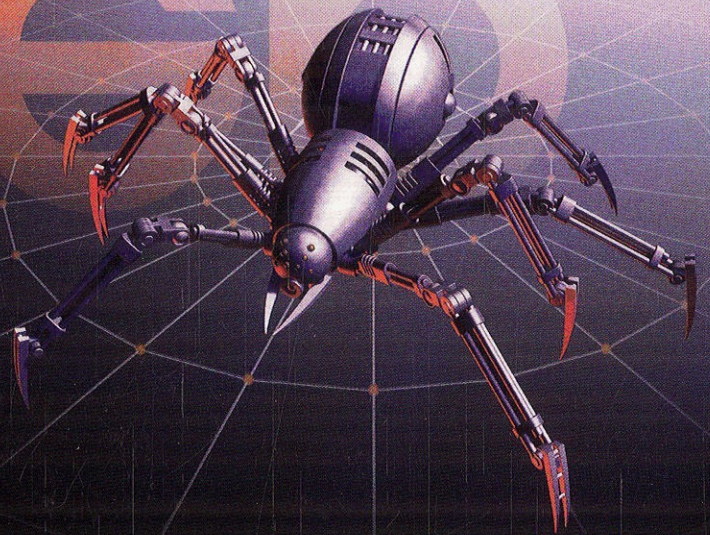


Second Edition

Web

APPLICATION ARCHITECTURE



Principles, Protocols
and Practices

Leon Shklar and Rich Rosen

Contents

About the Authors	xv
Preface	xvi
Acknowledgments	xvii
1 Introduction	1
1.1 History and Pre-History of the Web	1
1.2 From Web Pages to Web Sites	3
1.3 From Web Sites to Web Applications	4
1.4 Web 2.0: On-line Communities and Collaboration	4
1.5 The Brave New World of AJAX	5
1.6 Focus of This Book	5
1.7 What Is Covered in This Book	7
1.8 Bibliography	9
2 Core Internet Protocols	11
2.1 Historical Perspective	12
2.2 TCP/IP Architecture	13
2.2.1 Protocol layers	13
2.2.2 Comparison with OSI model	14
2.2.3 The client–server paradigm	14
2.3 TCP/IP Application Services	16
2.3.1 Telnet	16
2.3.2 E-mail	17
2.3.3 Message forums	23
2.3.4 Chat and messaging protocols	24

2.3.5	Security and encryption	24
2.3.6	File server protocols	26
2.4	And Then Came the Web ...	27
	Questions and Exercises	28
2.5	Bibliography	28
3	Birth of the Web: HTTP	29
3.1	Historical Perspective	29
3.1.1	CERN: birthplace of the web	29
3.1.2	Building blocks of the web	30
3.2	Uniform Resource Locator	30
3.3	Fundamentals of HTTP	32
3.3.1	Request–response paradigm	33
3.3.2	Stateless protocol	33
3.3.3	Structure of HTTP messages	34
3.3.4	Request methods	36
3.3.5	Status codes	41
3.4	Better Information Through Headers	44
3.4.1	Support for content types	46
3.4.2	Caching control	49
3.4.3	Security	51
3.4.4	Session support	53
3.5	Evolution of the HTTP Protocol	56
3.5.1	Virtual hosting	57
3.5.2	Caching support	58
3.5.3	Persistent connections	59
3.6	Summary	60
	Questions and Exercises	60
3.7	Bibliography	61
4	HTML and Its Roots	63
4.1	Standard Generalized Markup Language	64
4.1.1	SGML declaration	66
4.1.2	Document Type Definition	68
4.2	HTML	72
4.2.1	Evolution of HTML	72
4.2.2	Structure and syntax	73
4.3	HTML Rendering	79
4.3.1	Cascading Style Sheets	79
4.3.2	Associating styles with HTML documents	80
4.4	Summary	82
	Questions and Exercises	82
4.5	Bibliography	83

5	XML Languages and Applications	85
5.1	Core XML	86
5.1.1	XML documents	87
5.1.2	XML DTD	88
5.1.3	XML Schema	91
5.2	XHTML	95
5.2.1	HTML 5	97
5.2.2	XHTML MP	97
5.3	Web Services	100
5.3.1	SOAP	100
5.3.2	Representational State Transfer (REST)	105
5.4	XSL	108
5.4.1	XSLT	109
5.4.2	XSL Formatting Objects	111
5.4.3	What is so important about XSL?	116
5.5	Summary	118
	Questions and Exercises	119
5.6	Bibliography	120
5.7	Web Links	120
5.8	Endnotes	120
6	Web Servers	121
6.1	Basic Operation	122
6.1.1	HTTP request processing	123
6.1.2	Delivery of static content	125
6.1.3	Delivery of dynamic content	127
6.2	Mechanisms for Dynamic Content Delivery	135
6.2.1	Beyond CGI and SSI	135
6.2.2	Native APIs (ISAPI and Apache Server API)	135
6.2.3	FastCGI	135
6.2.4	Template processing	136
6.2.5	Servlets	137
6.2.6	Java Server Pages	138
6.2.7	Future directions	139
6.3	Advanced Functionality	140
6.3.1	Virtual hosting	140
6.3.2	Chunked transfers	141
6.3.3	Caching support	142
6.3.4	Extensibility	143
6.4	Server Configuration	143
6.4.1	Directory structure	143
6.4.2	Execution	144
6.4.3	Address resolution	145

6.4.4	MIME support	146
6.4.5	Server extensions	146
6.5	Server Security	147
6.5.1	Securing the installation	147
6.5.2	Dangerous practices	148
6.5.3	Secure HTTP	149
6.5.4	Firewall configurations	149
6.5.5	HTTP proxies	150
6.6	Summary	150
	Questions and Exercises	151
6.7	Bibliography	152
7	Web Browsers	153
7.1	Overview of Browser Functionality	154
7.2	Architectural Considerations	155
7.3	Overview of Processing Flow in a Browser	157
7.3.1	Transmitting a request	157
7.3.2	Receiving a response	159
7.4	Processing HTTP Requests	162
7.4.1	Constructing the request line	163
7.4.2	Constructing the headers	165
7.4.3	Constructing the request body	166
7.4.4	Transmitting the request	167
7.5	Processing HTTP Responses	167
7.5.1	Processing successful responses	168
7.5.2	Processing responses with other status codes	170
7.6	Cookie Coordination	172
7.7	Privacy and P3P	173
7.8	Complex HTTP Interactions	174
7.8.1	Caching	174
7.8.2	Authorization: challenge and response	178
7.8.3	Using common mechanisms for data persistence	179
7.8.4	Requesting supporting data items	180
7.8.5	Multimedia support: helpers and plug-ins	182
7.9	Summary	184
	Questions and Exercises	186
7.10	Bibliography	187
7.11	Web Links	188
7.12	Endnotes	188
8	Active Browser Pages: From JavaScript to AJAX	189
8.1	Pre-History	191
8.2	JavaScript	191

8.2.1	Manipulating page content	194
8.2.2	Client-side form validation	196
8.2.3	Hovering behaviors: image rollover	199
8.2.4	JavaScript Object Notation	201
8.2.5	Summary	202
8.3	Cascading Style Sheets	203
8.3.1	Format of CSS rules	204
8.3.2	Hovering behaviors: the a:hover pseudo-class	208
8.3.3	Summary	209
8.4	DHTML	209
8.4.1	Inner workings	210
8.4.2	Controlling content visibility	210
8.4.3	Leveraging toolkits	212
8.4.4	Client-side validation using toolkits	216
8.4.5	Hovering behaviors using toolkits	217
8.4.6	Widgets	219
8.4.7	Summary	221
8.5	AJAX	221
8.5.1	Content injection: manual approach	222
8.5.2	Content injection: using toolkits	224
8.5.3	Auto-completion	225
8.5.4	Remote validation	228
8.5.5	Where does DHTML end and AJAX begin?	229
8.5.6	Summary	231
8.6	Case Study: 5-Star Rating	231
8.6.1	Designing a star-rating component	232
8.6.2	When you click upon a star: what happens on the server?	236
8.7	Summary	236
	Questions and Exercises	237
8.8	Bibliography	237
8.9	Web Links	238
8.10	Endnotes	238
9	Approaches to Web Application Development	239
9.1	Taxonomy of Web Application Approaches and Frameworks	240
9.1.1	Programmatic approaches	240
9.1.2	Template approaches	241
9.1.3	Hybrid approaches	241
9.1.4	Frameworks	242
9.2	Comparative Survey of Web Application Approaches and Frameworks	245
9.2.1	CGI and FastCGI	245
9.2.2	Server-Side Includes (SSI)	246
9.2.3	PHP	246

9.2.4	Java Servlet API	247
9.2.5	Cold Fusion	249
9.2.6	Velocity	251
9.2.7	Active Server Pages and .NET	252
9.2.8	Java Server Pages	254
9.2.9	JSP Model 2	257
9.2.10	Java Standard Tag Library	258
9.2.11	Struts	260
9.2.12	Java Server Faces	261
9.2.13	JBoss Seam	264
9.2.14	Rapid application development: Ruby on Rails	264
9.3	Summary	267
	Questions and Exercises	272
9.4	Bibliography	272
10	Web Application Primer 1: Struts and JSTL	275
10.1	Case Study: Virtual Realty Listing Services	276
10.2	Application Requirements	278
10.3	Technology Choices	279
10.4	Overview of Struts	280
10.5	Structure of the VRLS Application	283
10.5.1	Configuration	284
10.5.2	Controller components	287
10.5.3	View components	291
10.5.4	Model components	297
10.6	Design Decisions	299
10.6.1	Abstracting functionality into service classes	299
10.6.2	Including embedded pages to support co-branding	301
10.6.3	Creating and modifying customer profiles in one task	302
10.7	Suggested Enhancements	303
10.7.1	Adding an administrative interface	303
10.7.2	Enhancing the signup process through e-mail authentication	304
10.7.3	Improving partner recognition through a persistent cookie	305
10.7.4	Adding caching functionality to the DomainService Class	306
10.7.5	Paging through cached search results	307
10.7.6	Using XML and XSLT for view presentation	308
10.7.7	Tracking user behavior	310
10.7.8	Using an object-relational mapping tool	310
10.7.9	Adding DHTML and AJAX for an enhanced user experience	311
10.8	Summary	312
	Questions and Exercises	312
10.9	Bibliography	313

11 Web Application Primer 2: Ruby on Rails	315
11.1 Comparing Rails with Java EE	316
11.1.1 Similarities	316
11.1.2 Differences	317
11.2 Application Requirements	317
11.3 Building the Administrative Interface as a Rails Application	318
11.3.1 Downloading and installing Ruby and Rails	318
11.3.2 Building an application skeleton	318
11.3.3 Creating a new project and configuring the database	321
11.3.4 Scaffolding for the model, view, and controller classes	322
11.3.5 Enhancing the application	326
11.4 Benefits and Drawbacks of Using Rails	339
11.4.1 How rapid is rapid application development?	339
11.4.2 Database support	340
11.4.3 Limitations of scaffolding	340
11.4.4 Scalability	341
11.4.5 Performance and clustering	341
11.4.6 Version 2.0 issues	341
11.4.7 Is Rails web-designer-friendly?	342
11.5 Whither Enterprise Java?	343
11.6 Summary	344
Questions and Exercises	344
11.7 Bibliography	345
11.8 Web Links	345
11.9 Endnotes	345
12 Search Technologies	347
12.1 Overview of Algorithms	348
12.1.1 Historical perspective	348
12.1.2 Basic vector-space algorithm	349
12.1.3 Common enhancements	355
12.1.4 Word clustering	355
12.1.5 Custom biases	357
12.2 Searching the Web	358
12.2.1 Google page ranking	358
12.2.2 Web spiders	359
12.2.3 Summary	361
12.3 Site Search Applications	361
12.3.1 General architecture	362
12.3.2 Lucene	363
12.3.3 OneBox applications	367

12.4	Search Engine Optimization	369
12.4.1	Robots.txt	370
12.4.2	Sitemaps	372
12.4.3	Sitemap extensions	376
12.4.4	Site and URL structure	376
12.4.5	Black Hat SEO	377
12.5	Summary	378
	Questions and Exercises	379
12.6	Bibliography	380
12.7	Web Links	380
12.8	Endnotes	380
13	Trends and Directions	381
13.1	XML Query Language	382
13.2	Semantic Web	385
13.2.1	Resource Description Framework (RDF)	386
13.2.2	Composite Capabilities/Preference Profiles	392
13.2.3	RDF query language	393
13.3	Future of Web Application Frameworks	396
13.3.1	One more time: separate content from presentation	396
13.3.2	Use the right tools for the job	398
13.3.3	Simplicity	399
13.4	Current Trends	404
13.4.1	Everything old is new again	404
13.4.2	Social networking and community web sites	405
13.4.3	Cloud computing and “Weblications”	405
13.5	Summary	407
	Questions and Exercises	407
13.6	Bibliography	408
13.7	Web Links	408
13.8	Endnotes	408
14	Conclusions	409
	Index	413