Apache Overview HOWTO

Daniel Lopez Ridruejo, `ridruejo@apache.org` v0.9, 10 ottobre 2002

Questo documento passa in rassegna i diversi progetti che riguardano Apache, come il server Apache HTTP ed il motore per Servlet e JSP Tomcat; inoltre fornisce indicazioni per ulteriori approfondimenti. Traduzione a cura di Tito Rizzo, `tito.rizzo@tiscali.ite` revisione a cura di Giuseppe Briotti, `g.briotti@mclink.it`.

Indice

1 Introduzione

Questo documento fornisce un'introduzione al mondo di Apache, includendo i vari progetti della Apache Software Foundation come il web server Apache e i vari software di terze parti, sia commerciali che open source. Apache è *il server più popolare di Internet* `http://www.netcraft.com/survey/` . I nuovi utilizzatori di Apache, specie quelli che provengono da Windows, sono spesso inconsapevoli delle potenzialità di Apache, della sua estensibilità e, più in generale, della efficiente interrelazione tra i vari componenti. Lo scopo di questo documento è quello di fornire una panoramica di tali possibilità, con una breve descrizione di ciascuna e indicazioni per ulteriori approfondimenti. Le informazioni sono state raccolte da varie fonti, includendo le pagine web dei vari progetti, interventi a conferenze, mailing list e siti web di Apache, oltre alla mia personale esperienza. Tali autori meritano la più ampia considerazione, poiché senza il loro lavoro, questo documento non sarebbe stato né possibile né necessario.

1.1 Apache Software Foundation

La Apache Software Foundation fornisce il supporto alla comunità che si occupa dei progetti di software open-source per Apache, progetti caratterizzati da un processo di sviluppo collaborativo e basato sul consenso, da una licemza aperta e pragmatica e dal desiderio di creare software di alta qualità, che apra nuove strade *in questo campo. Ci reputiamo non semplicemente un gruppo di progetti che abbiano in comune lo stesso server, ma piuttosto una comunità di sviluppatori e di utenti.*

La ASF è la casa madre di innumerevoli progetti Open Source di successo, come il motore Tomcat per Servlet/JSP e lo strumento di compilazione ANT.

Per saperne di più sulla fondazione si veda *qui* `http://www.apache.org/foundation/` .

1.2 Struttura del documento

La prima parte del documento si occupa del Web Server Apache e dei relativi moduli. Tratta la storia, l'architettura e le possibilità di questo server e descrive in quali modi lo si possa estendere e personalizzare.

La seconda parte del documento tratta i progetti della Apache Software Foundation, come quelli che costituiscono le comunità Jakarta e Java XML. Piuttosto che organizzare i progetti in base al linguaggio di programmazione o alla tecnologia impiegata, si è preferito organizzarli in base alle funzionalità che offrono.

2 Apache

Apache è il principale server web, con una quota di mercato superiore al 60%, secondo *Netcraft survey* `http://www.netcraft.com/survey` . Al successo di Apache hanno contribuito svariati fattori chiave:

- La *licenza di Apache* `http://www.apache.org/LICENSE.txt` . è una licenza open source, di tipo BSD, che consente l'utilizzo di Apache per scopi sia commerciali che non.
- Una comunità di abili *sviluppatori* `http://www.apache.org/foundation/members.html`

  con una varietà di bagagli culturali e un processo di sviluppo aperto basato su meriti tecnici.
- Architettura modulare. Gli utenti di Apache possono facilmente aggiungere nuove funzionalità o adattare Apache al loro specifico ambiente di lavoro.
- Portabilità: Apache funziona praticamente su tutte le versioni di Unix (e di Linux), Windows, BeOs, sui mainframe...
- Robustezza e sicurezza.

Svariate società commerciali hanno adottato soluzioni basate su Apache per i loro prodotti, incluse *Oracle* `http://www.oracle.com` , *Red Hat* `http://www.redhat.com` e *IBM* `http://www.ibm.com` . In aggiunta, *Covalent* `http://www.covalent.net` fornisce moduli aggiuntivi e supporto 24x7 per Apache.

I seguenti siti web usano Apache o uno dei suoi derivati. Se Apache va bene per loro, dovrebbe andare bene anche per voi :)

- *Amazon.com* `http://www.amazon.com`
- *Yahoo!* `http://www.yahoo.com`
- *W3 Consortium* `http://www.w3c.org`
- *Financial Times* `http://www.ft.com`
- *Apple* `http://www.apple.com`
- *MP3.com* `http://www.mp3.com`
- *Stanford* `http://www.stanford.edu`

Dal *sito web di Apache* `http://httpd.apache.org` :

Il progetto del Server HTTP Apache rappresenta uno sforzo per sviluppare e mantenere un server HTTP open-source per i moderni sistemi operativi, incluso Unix e Windows NT. Il fine di questo progetto è quello di fornire un server sicuro, efficiente e estensibile che garantisca servizi HTTP in aderenza con gli standard HTTP correnti.

Il server Apache iniziò la sua vita come modifica del server Web NCSA, uno dei primi server HTTP. Per saperne di più sulla storia di Apache si veda *qui* `http://httpd.apache.org/ABOUT_APACHE.html` .

Il progetto Apache è poi cresciuto oltre la semplice implementazione di un server web, occupandosi dello sviluppo di altre tecnologie critiche dal lato server. La Apache Software Foundation, che sarà descritta più oltre, serve da ala protettiva per questi progetti.

2.1 Architettura

Ci sono due versioni principali di Apache, la serie 1.3 e la serie 2.0. Per quanto siano da considerarsi entrambe prodotti di qualità, nondimeno differiscono per architettura e possibilità.

2.1.1 2.1.1 Apache 1.3

Apache 1.3 è stato portato su una grande varietà di piattaforme Unix, ed è il server Web più largamente utilizzato in Internet.

Server Web basato su processi Apache 1.3 su Unix è un server Web basato su processi. Il programma Apache, al suo avvio, genera (fork) svariati processi figli; con il fork un processo primario genera copie identiche di se stesso, chiamate *figli*. Ognuno di tali figli può servire una richiesta indipendente dalle altre, con il vantaggio di migliorare la stabilità: se uno di tali figli ha un comportamento anomalo (va fuori controllo o ha perdite di memoria) può essere interrotto senza alcun effetto sugli altri. La stabilità è conseguita a spese delle prestazioni. Nella maggior parte dei sistemi Unix, la creazione di processi e il cambiamento del contesto (assegnazione di tempo del processore a ogni processo) sono operazioni costose in termini di risorse di sistema, dal momento che i processi sono isolati gli uni dagli altri e non possono quindi facilmente condividere codice e dati.

Supporto per Windows Apache 1.3 è la prima versione di Apache a supportare Windows, per quanto tale adattamento non sia considerato stabile come quello su Unix, visto che il server è stato progettato prendendo come riferimento Unix, mentre la versione per Windows ha costituito una aggiunta successiva, non molto bene integrata.

Modularità Apache 1.3 possiede un'architettura modulare, che permette di abilitare o disabilitare moduli per aggiungere o rimuovere funzionalità al server Web. È possibile personalizzare Apache per migliorarne le prestazioni e la sicurezza. In aggiunta ai moduli di base, esiste un gran numero di moduli di terze parti, con i quali è possibile ampliare ulteriormente le funzionalità del server.

2.1.2 2.1.2 Apache 2.0

Apache 2.0 è la più recente e migliore versione del server Apache, la cui architettura contiene significativi miglioramenti rispetto alla serie 1.3, tra i quali si possono menzionare i seguenti.

Moduli multiprocesso Apache 2.0 astrae l'architettura di elaborazione delle richieste tramite moduli speciali, chiamati Multi Processing modules (MPMs), ossia moduli Multi Processo, grazie ai quali Apache può essere configurato come un server basato su processi, come un server puramente basato su thread o come un misto di tali modelli. I thread sono inglobati all'interno dei processi e sono eseguiti simultaneamente; a differenza dei processi, possono condividere sia dati che codice, risultando così più leggeri, tanto che in molti casi i server di tipo threaded hanno una migliore scalabilità rispetto ai server basati su processi. Il server risulta, per contro, meno affidabile, perché se un thread ha un comportamento anomalo può causare la corruzione di codice o di dati appartenenti ad altri thread.

Moduli di protocollo La gestione dei protocolli in Apache 2.0 è stata incapsulata in un suo proprio livello, rendendo possibile la scrittura di moduli per protocolli diversi da HTTP, quali POP3 per la posta o FTP per il trasferimento di file. Questi moduli di protocollo possono trarre vantaggio dalla solida ossatura del server e dalla funzionalità dei moduli, come l'autenticazione e la generazione di contenuti dinamici. Questo significa che, per esempio, è possibile autenticare i propri utenti POP3 tramite lo stesso database di utenti che Apache usa per le richieste web e che il contenuto FTP può essere generato dinamicamente grazie a PHP, CGI o qualunque altra delle tecnologie descritte più avanti in questo stesso documento.

Architettura dei moduli e dei filtri Apache 2.0 conserva l'architettura modulare di 1.3 con l'aggiunta di un ulteriore meccanismo di estensione rappresentato dai filtri, che consentono ai moduli di modificare il contenuto generato da altri moduli. Possono criptare, effettuare la scansione alla ricerca di virus o comprimere non solo i file statici, ma anche i contenuti generati dinamicamente.

Problemi di compatibilità Sfortunatamente, per quanto il modulo API sia simile in entrambe le versioni, non è comunque identico e i moduli di Apache 1.3 necessitano di essere adattati alla nuova architettura. La maggioranza dei moduli più importanti, come PHP o mod_perl hanno già la versione per Apache 2.0, mentre altri, come mod_dav e mod_ssl, sono ora parte della distribuzione del server. Il funzionamento dei moduli su un'architettura threaded richiede modifiche specifiche ai moduli. I moduli distribuiti con Apache hanno beneficiato di tali modifiche e sono ora considerati 'thread-safe', ma i moduli e le librerie di terze parti possono non esserlo, per cui, se si ha bisogno di qualcuno di questi ci si deve limitare a usare Apache come server basato unicamente su processi.

Portabilità Apache funziona ugualmente bene sia su piattaforme Windows che Unix grazie alla libreria Apache Portable Runtime (APR), che astrae le differenze tra sistemi operativi, come le API di accesso al network o ai file. Adattare Apache ad una nuova piattaforma risulta spesso altrettanto semplice che adattare il livello di astrazione rappresentato dalla Apache Portable Runtime, la quale provvede anche alla messa a punto e all'ottimizzazione specifica per una determinata piattaforma.

2.2 Sicurezza

Apache fornisce svariati moduli orientati alla sicurezza per regolamentare gli accessi al server.

2.2.1 Autenticazione

I moduli di autenticazione permettono di determinare l'identità di un client, solitamente attraverso la verifica del nome utente e della password rispetto ai dati registrati in un database. Apache comprende inoltre moduli di autenticazione che connettono Apache con strutture o database di sicurezza, includendo tra gli altri: NT Domain controller, Oracle, mySQL, PostgresSQL e così via.

Particolarmente interessanti sono i moduli LDAP, in quanto consentono l'integrazione con servizi di directory a livello di compagnia o di impresa. È possibile trovare tali moduli presso . Un modulo LDAP per Apache 2.0 può essere scaricato *dal sito web di Apache* `http://httpd.apache.org/docs-2.0/mod/mod_auth_ldap.html` .

2.2.2 Controllo degli accessi

Apache fornisce il modulo mod_access che può limitare l'accesso alle risorse basandosi su parametri della richiesta del client, come la presenza di una specifica intestazione o dell'indirizzo IP o del nome host del client. Moduli di terze parti permettono di limitare l'accesso per quei client responsabili di comportamenti irregolari, come spiegato più oltre nelle sezioni dedicate alle prestazioni e al controllo dell'ampiezza di banda.

2.2.3 SSL/TLS

I protocolli Secure Sockets Layer/Transport Layer Security consentono di criptare i dati trasmessi tra il server Web e il client. In Apache 1.3, i protocolli sono implementati con il mod_ssl, distribuito separatamente dal *sito web di mod_ssl* `http://www.modssl.org` , che rende necessario applicare alcune patch al server, a causa della regolamentazione sull'esportazione di software di criptatura. La maggior parte di tali restrizioni sono nel frattempo decadute e, a partire da Apache 2.0, il modulo mod_ssl è incluso in Apache come modulo base.

2.3 Proxy

Un proxy è un programma che effettua richieste per conto di qualcun altro. Esistono diversi tipi di proxy per il Web. Un tradizionale proxy HTTP, chiamato anche *forward proxy*, accetta richieste dai client (solitamente browser Web), contatta il server remoto e restituisce le risposte.

Un reverse proxy è un server Web collocato di fronte a altri server, per i quali costituisce un accesso unificato sgravandoli da determinate incombenze, come l'elaborazione SSL.

Apache supporta entrambi i tipi di proxy, l'immagazzinamento di una cache di contenuti Web e differenti proxy di backend tipo FTP.

2.4 Prestazioni e scalabilità

La grezza prestazione è solo uno dei fattori da considerare in un server Web (di solito flessibilità e stabilità vengono prima).

Detto questo, ci sono soluzioni per migliorare le prestazioni in server web che forniscono contenuti statici e sottoposti a carichi gravosi. Per le imprese di hosting, Apache fornisce anche sistemi per misurare e controllare l'uso dell'ampiezza di banda. In questo contesto, abitualmente la regolazione si traduce in un rallentamento della distribuzione del contenuto in base al file richiesto, ad uno specifico indirizzo IP del client e così via. Questo è fatto al fine di prevenire gli abusi.

- **mod_mmap**: Incluso nella distribuzione corrente di Apache 1.3, traccia in memoria una lista staticamente configurata di file richiesti spesso, ma modificati non frequentemente. In Apache 2 questa funzionalità è inclusa nel mod_file_cache.
- ***Mod_ bandwidth*** `http://www.cohprog.com/mod_bandwidth.html` : Questo modulo di Apache 1.3 abilita l'impostazione di limiti di ampiezza di banda per l'intero server o per singole connessioni, in base ad una specifica directory, alla dimensione dei file e al dominio IP remoto.

- ***Modulo di condivisione dell'ampiezza di banda*** `http://www.topology.org/src/bwshare/README.html` : fornisce la regolazione ed il bilanciamento dell'ampiezza di banda in base all'indirizzo IP del client. Supporta Apache 1.3 e le prime versioni di Apache 2.
- ***Mod_ throttle*** `http://www.snert.com/Software/Throttle/index.shtml` :Regola l'ampiezza di banda per host virtuale o per utente. Per Apache 1.3.

2.4.1 Bilanciamento del Carico

Usando il reverse proxy di Apache e il mod_rewrite è possibile ottenere un processo di distribuzione delle richieste tra una varietà di server web di backend. Per maggiori informazioni si veda

In aggiunta, il mod_backhand è un modulo di Apache 1.3 che consente una redirezione di richieste HTTP priva di giunzioni, tra un server web e l'altro. Questa redirezione può essere usata per coinvolgere macchine le cui risorse siano sottoutilizzate, ottenendo così un granulare bilanciamento del carico delle richieste web. Per maggiori informazioni si veda .

2.4.2 Compressione

Apache 2.0 include mod_deflate, un modulo di filtraggio che comprime il contenuto prima di inviarlo al client. Questo limita lo spreco di banda ma può avere un impatto negativo sulle prestazioni. Il *modulo mod_ gzip* `http://www.remotecommunications.com/apache/mod_gzip/`

fornisce questa funzionalità per Apache 1.3.

2.5 Script CGI

CGI sta per Common Gateway Interface. I programmi CGI sono programmi esterni che vengono invocati quando un utente richiede una certa pagina. Il programma CGI riceve le informazioni dal server web (valori variabili del form, tipo di browser, indirizzo IP del client e così via) e usa tali informazioni per produrre una pagina web per il client.

Apache ha il supporto per CGI e c'è un modulo di terze parti per Apache 1.3 che garantisce il supporto per il protocollo FastCGI, il quale evita le penalizzazioni prestazionali associate all'avvio ed all'arresto di un programma CGI a ogni richiesta. Lo si può trovare a

2.6 Integrazione delle Piattaforme di Sviluppo

Le applicazioni web sono scritte in linguaggi di alto livello come Java, Perl, C# e così via e Apache dispone di svariati moduli che ne permettono l'integrazione con il server. In molti casi i moduli espongono la API di Apache, così che tutti i moduli Apache possano essere scritti in questi linguaggi.

2.6.1 Perl

mod_perl `http://perl.apache.org/` è un veterano tra i progetti Apache e uno dei più fortunati. Ingloba all'interno di Apache un interprete Perl e consente l'accesso, tramite Perl, all'interno di Apache. Permette che interi moduli possano essere scritti in Perl o in un misto di Perl e di codice C. Nella versione 1.3 di Apache, in ogni processo figlio deve essere incluso un interprete, visto che il server si basa sul multiprocesso. In siti dinamici a alto traffico, l'accresciuta dimensione può fare la differenza. Nelle versioni threaded di Apache 2.0, il mod_perl permette la condivisione del codice, dei dati e dello stato della sessione tra i vari interpreti, tendendo a una soluzione più snella e veloce.

mod_perl costituisce di per sé un'altra piattaforma, con una grande varietà di moduli disponibili, come *Mason* `www.masonhq.com` e

Embperl `http://perl.apache.org/embperl/` per l'inclusione del Perl in pagine HTML e *AxKit* `axkit.org` per template ricavati da XML.

2.6.2 PHP

Dal sito web di *PHP* `http://www.php.net` : *PHP è un linguaggio di scripting, lato server, indipendente dalla piattaforma, inserito in HTML.*. È il

modulo per Apache più popolare `http://www.securityspace.com/s_survey/data/man.200209/apachemods.html`

e questo grazie ad una varietà di ragioni:

- La curva di apprendimento è piuttosto bassa
- Vasta documentazione
- Ampio supporto per database
- Modularità

PHP ha una concezione modulare. Tra molti altri, ci sono moduli che garantiscono il supporto per:

- Connettività con database per i database più diffusi, come Oracle, server MS-SQL, interfaccia ODBC, MySQL, mSQL, PostgreSQL e via dicendo.
- supporto XML
- Trasferimento di file: FTP
- HTTP
- Supporto per directory: LDAP
- Supporto per la posta: IMAP, POP3, NNTP
- Generazione di documentazione PDF
- CORBA
- SNMP

Non bisogna fare altro che compilare/usare il modulo di cui si ha bisogno. PHP può essere usato con Apache, come CGI esterno o con altri server web. È indipendente dalla piattaforma e funziona sulla maggior parte delle versioni di Unix e di Windows. Chiunque provenga da un ambiente Windows, avrà probabilmente usato Internet Information Server con Active Server Pages e MS-SQL Server. Un sostituto comune per questo terzetto nel mondo Unix è Apache con PHP e MySQL. Dal momento che PHP funziona:

- con Apache e con Microsoft IIS
- con MySQL e con il server MS-SQL
- su Unix e su Windows

si ha a disposizione un percorso di migrazione piacevole e graduale da una soluzione incentrata su Microsoft a soluzioni basate su Unix.

2.6.3 Python

Python è un popolare linguaggio di scripting orientato agli oggetti.

Mod_Python `http://www.modpython.org` , che è ora un progetto Apache ufficiale, permette di integrare Python con il server web Apache. È possibile sviluppare applicazioni web complesse o sveltire script CGI Python esistenti. Le versioni recenti funzionano su Apache 2.0.

2.6.4 Tcl

Il *Progetto Tcl Apache* `http://tcl.apache.org` integra Tcl con il server web Apache. Tcl è un linguaggio di scripting leggero e estensibile. Per saperne di più su Tcl si veda *qui* `http://tcl.activestate.com/` . Ci sono attualmente svariati moduli sotto l'ala di Apache Tcl:

- Sia *Mod_dtcl* `http://tcl.apache.org/mod_dtcl/` che

  Neowebscript `http://tcl.apache.org/neowebscript/` consentono di includere Tcl nelle pagine HTML. *Rivet* `http://tcl.apache.org/rivet/` combina il meglio di entrambi i moduli.
- Il *Mod_tcl* `http://tcl.apache.org/mod_tcl/mod_tcl.html` sceglie un approccio simile a mod_perl, esponendo la API di Apache.
- *WebSH* `http://tcl.apache.org/websh/` fornisce un ambiente Tcl per applicazioni Web.

2.6.5 Tecnologie Microsoft

Svariati moduli permettono l'integrazione con i linguaggi e le tecnologie Microsoft quali la struttura .Net o le Active Server Pages.

.Net *mod_haydn* `http://haydn.sourceforge.net/` integra

Mono `http://www.go-mono.com` con Apache e espone la API di Apache alla struttura .Net, permettendo, per esempio, la scrittura di moduli in C#.

Covalent `http://www.covalent.net` fornisce mod_asp.net, un modulo commerciale Windows che permette ad Apache di lanciare applicazioni ASP.Net, consentendo di rimpiazzare Microsoft IIS.

ASP ASP, che sta per Active Server Pages, è una tecnologia Microsoft che permette di inserire codice, abitualmente Visual Basic, nelle pagine HTML. Diverse compagnie come *ChilliSoft* `http://www.chilisoft.com/` e *Stryon* `http://www.stryon.com/` forniscono prodotti in grado di far funzionare applicazioni ASP in ambiente Unix.

ISAPI ISAPI è una API utilizzabile per estendere Microsoft IIS, in modo simile a come si userebbe la API di Apache. Apache include un modulo mod_isapi che rispecchia questa funzionalità e permette di usare moduli ISAPI.

2.6.6 Java

La maggior parte dei server di applicazioni, come quelli di Oracle, IBM e BEA forniscono moduli per l'integrazione con il server web Apache. In aggiunta, svariati moduli come mod_jk e mod_webapp permettono la connessione con Tomcat, un contenitore di Servlet e JSP (JavaServer Pages) che fa sempre parte della Apache Software Foundation.

2.6.7 Moduli per altri linguaggi

Questo documento ha descritto moduli per linguaggi lato server popolari come Perl, Python e PHP. Moduli aggiuntivi per altri linguaggi (JavaScript, Haskell, Ruby e altri) si possono trovare nella *directory dei moduli di Apache* `http://modules.apache.org` .

2.7 Gestione

Una parte importante dell'amministrazione di server Web comprende la compilazione, la configurazione e il controllo di diversi server.

2.7.1 Strumenti di compilazione

Apache può essere esteso e personalizzato in molti modi differenti. L'integrazione di moduli esterni con il server può talvolta costituire un compito difficile. Strumenti come *Apache Toolbox* `http://www.apachetoolbox.com` possono semplificare questo compito, fornendo un ambiente di compilazione dotato di menu.

2.7.2 Interfacce Utente per Apache

Apache è configurato tramite file di testo, il che talvolta può risultare ostico, specie per chi provenga da un ambiente Windows. Esistono strumenti grafici open source che facilitano questo compito:

- *Comanche* `http://www.comanche.org` , provato dal sottoscritto, indipendente dalla piattaforma, funziona su Unix/Linux, Windows e Mac.
- *Webmin* `http://www.webmin.com/webmin/` : Una interfaccia piacevole, basata sul web.
- *gui.apache.org* `http://gui.apache.org` : Interfacce grafiche per il progetto Apache. I programmi si trovano a vari livelli di sviluppo.

2.7.3 SNMP

SNMP sta per Simple Network Management Protocol. Consente il controllo e la gestione di server di rete, attrezzature e così via. I moduli SNMP per Apache aiutano nella gestione di vasti spiegamenti di server web, misurano la qualità del servizio offerto e l'integrazione di Apache con le preesistenti infrastrutture di gestione.

- *Mod SNMP* `http://www.simpleweb.org/software/packages/mod-snmp/` open source per Apache 1.3.
- *Covalent SNMP* `http://www.covalent.net` fornisce un modulo SNMP commerciale, supporto per i più recenti standard SNMPv3, integrazione con HP-Openview, Tivoli e via dicendo.

2.8 Publishing

Gli autori di contenuti Web richiedono uno strumento per gestire tali contenuti e per l'upload sul server. Uno dei protocolli usati a questo scopo è DAV (Distributed Authoring and Versioning). DAV è un'estensione del protocollo HTTP che permette a utenti e applicazioni di pubblicare e modificare contenuti Web. La tecnologia DAV è ampiamente implementata, Microsoft la supporta a livello di sistema operativo (WebFolders) e nella sua suite Office. Lo stesso vale per Apple OS-X e una varietà di prodotti di terze parti come Adobe, Oracle e così via. Il modulo mod_dav per Apache 1.3 si può trovare presso . In Apache 2.0, il mod_dav è incluso nella distribuzione di base.

Antecedentemente a DAV, Microsoft disponeva di un proprio protocollo per il publishing, integrato nell'applicazione Microsoft FrontPages. Un supporto lato server per FrontPages si può ottenere con i moduli di , tuttavia, a causa del modo con cui si integrano con Apache, non sono da ritenere sicuri.

2.9 Moduli di Protocollo

Apache 2.0 introduce il concetto di moduli di protocollo. Il che significa che gli sviluppatori possono riutilizzare l'infrastruttura del server Apache per implementare nuovi protocolli, come quelli che si occupano della posta o del trasferimento di file. mod_ftp è un modulo FTP commerciale basato su Apache fornito da *Covalent* `http://www.covalent.net` . *mod_pop3* `http://cvs.apache.org/viewcvs.cgi/httpd-pop3/` è un modulo open source che implementa il protocollo POP3, comunemente usato per recuperare i messaggi di posta dai relativi server.

2.10 Hosting virtuale

Apache fornisce supporto estensivo per l'hosting virtuale, il che realizza la possibilità di servire vari siti web con un singolo server. In Apache 2.0, con i MPM si possono avere svariati processi figli, ognuno dei quali al servizio di un dominio differente sotto identificativi di utente Unix differenti. Questo è molto importante per la sicurezza in scenari di hosting condiviso, in quanto permette di isolare i clienti l'uno dall'altro. I seguenti sono alcuni moduli addizionali, alternativi, per l'hosting virtuale.

- *mod_dynvhost* `http://funkcity.com/0101/`
- *mod_pweb* `http://www.joytec.de/mod_pweb.html`
- *mod_v2h* `http://www.fractal.net/mod_v2h.tm`

2.11 Supporto commerciale

Apache è il server web d'elezione per svariate entità commerciali, incluse le grandi imprese. Tali compagnie chiedono determinati requisiti alle tecnologie da adottare, specie a quella che costituirà il nucleo della loro strategia Internet, come i server Web. Tali requisiti comprendono prestazioni, stabilità, possibilità di gestione, supporto, servizi professionali e integrazione con i sistemi proprietari. Compagnie commerciali, come

IBM `http://www.ibm.com` , *Red Hat* `http://www.redhat.com` e *Covalent* `http://www.covalent.net` , forniscono i prodotti e i servizi necessari affinché Apache soddisfi le esigenze della clientela Enterprise.

In aggiunta, svariate altre compagnie ed OEM offrono Apache come server web incluso nei loro prodotti.

3 Progetti ASF

Per quanto Apache sia probabilmente il più popolare, la Apache Software Foundation è la casa madre di molti altri progetti. Questa sezione fornisce una rassegna dei più rilevanti, organizzati con criterio logico. La maggioranza di essi appartiene al progetto Jakarta od al progetto XML. Il progetto Jakarta ospita i progetti basati su Java mentre il progetto XML ospita, manco a farlo apposta, i progetti collegati all'XML.

3.1 Applicazioni e Ambienti

I seguenti sono applicazioni e ambienti di sviluppo che fanno parte di ASF.

3.1.1 3.1.1 Server

I seguenti sono alcuni progetti ASF per il lato server.

Tomcat Tomcat è la nave ammiraglia del progetto Jakarta. Costituisce l'implementazione ufficiale di riferimento per le tecnologie Servlet Java e JavaServer Pages.

Per saperne di più si consulti la *homepage di Tomcat* `http://jakarta.apache.org/tomcat/` .

JAMES (Java Apache Mail Enterprise Server) Complementare rispetto alle altre tecnologie Apache lato server, JAMES fornisce *un server 100% Java, concepito come completa e portabile soluzione per un motore di posta per l'impresa, basato sui protocolli aperti attualmente disponibili (SMTP, POP3, IMAP, HTTP)*

Maggiori informazioni si possono trovare *qui* `http://jakarta.apache.org/james/` .

Lucene Jakarta Lucene è un motore per la ricerca testuale, ad alte prestazioni, perfettamente attrezzato, scritto in Java e parte del progetto Jakarta. Per maggiori informazioni si veda

Jetspeed *Jetspeed* `http://jakarta.apache.org/jetspeed/`

è un portale web scritto in Java. Possiede una API modulare che consente l'aggregazione di differenti fonti di dati (XML, SMTP, iCalendar)

3.1.2 3.1.2 Gestione dei contenuti

I seguenti sono progetti relativi alla gestione dei contenuti

Slide Slide è un ambiente ad alto livello per la gestione dei contenuti. Concettualmente, fornisce un'organizzazione gerarchica del contenuto binario che può essere immagazzinato in archivi dati arbitrari, eterogenei e distribuiti. In aggiunta, Slide integra servizi di sicurezza, di bloccaggio e di versioning. Provvede anche all'implementazione server e client di *WebDAV* `http://www.webdav.org` . Maggiori informazioni sulla *home page di Slide* `http://jakarta.apache.org/slide/index.html`

Alexandria Alexandria è un sistema integrato di gestione della documentazione. Mette insieme tecnologie comuni a molti altri progetti open source come CVS e JavaDoc. L'intento è quello di integrare codice sorgente e documentazione per incoraggiare la documentazione e la condivisione del codice. Maggiori informazioni presso

3.1.3 3.1.3 Ambienti

I seguenti sono infrastrutture per lo sviluppo per applicazioni

Turbine Turbine è un ambiente basato su servlet che permette a esperti sviluppatori Java di sviluppare rapidamente applicazioni web sicure. Turbine mette insieme una piattaforma per eseguire codice Java e componenti riutilizzabili. Alcune delle sue caratteristiche includono: integrazione con sistemi di template, sviluppo secondo il modello MVC, Access Control List, supporto per la localizzazione e così via. Maggiori informazioni si possono trovare sul *sito web di Turbine* `http://java.apache.org/turbine` .

Avalon Per chi abbia familiarità con sistemi Perl o BSD, Avalon è grosso modo l'equivalente di *CPAN* `http://www.cpan.org` o della collezione di Port per le tecnologie Java di Apache. Non si limita a fornire linee guida per un deposito comune di codice, ma va oltre: *costituisce uno sforzo per creare, progettare, costruire e mantenere una infrastruttura comune per applicazioni server scritte usando il linguaggio Java.* Fornisce i mezzi affinché progetti lato server in Java possano essere facilmente integrati e compilati l'uno sull'altro. Ulteriori informazioni si trovano sul *sito web di Avalon* `http://java.apache.org/avalon/` .

3.2 Presentazione

I seguenti sono sistemi di template, motori di trasformazione e altri progetti relativi a presentazioni.

3.2.1 Cocoon

Cocoon trae vantaggio da altre tecnologie XML di Apache come Xerces, Xalan e FOP per fornire un completo ambiente XML per il publishing. L'ambiente può dialogare con molte fonti di dati diverse e può rendere il contenuto in svariati formati finali come PDF, HTML, XML e RTF. Può essere eseguito come servlet o come programma da linea di comando. È possibile saperne di più, riguardo a Cocoon, sulla *homepage del progetto* `http://xml.apache.org/cocoon/`

3.2.2 Velocity

Velocity è un motore di template basato su Java. Può essere usato come utilità indipendente per generare codice sorgente, HTML, relazioni o può essere combinato con altri sistemi per fornire servizi di template. Velocity è basato sul paradigma Model View Controller, che rafforza la separazione tra il codice Java e il template HTML. È possibile saperne di più su Velocity *qui* `http://jakarta.apache.org/velocity/index.html` .

3.2.3 AxKit

AxKit `http://axkit.org` è un popolare Application Server basato su XML per mod_perl e Apache. Consente la separazione tra contenuto e presentazione e permette la conversione al volo da XML a ogni altro formato.

3.2.4 Xalan

Xalan è un processore XSLT disponibile per Java e C++. XSL è un linguaggio stylesheet per XML. La T sta per Transformation. XML è valido per l'archiviazione di dati strutturati (informazioni). Talvolta può rendersi necessario visualizzare tali dati o applicare qualche altra trasformazione. Xalan prende il documento XML originale, legge la configurazione per le conversioni (stylesheet) e lo restituisce come HTML, testo semplice o come altro documento XML. Per saperne di più su Xalan si consultino le homepage dei progetti *Xalan Java* `http://xml.apache.org/xalan-j/index.html` e *Xalan C++* `http://xml.apache.org/xalan-c/index.html` .

3.2.5 FOP

Dal sito web: *FOP è un'applicazione Java che legge l'albero degli oggetti di formattazione e lo trasforma in un documento PDF* . FOP prende dunque un documento XML e lo restituisce come PDF, in modo simile a quanto fa Xalan con l'HTML o i file di testo. Per saperne di più su FOP si veda *qui* `http://xml.apache.org/fop` .

3.3 Parser e Librerie per l'Accesso ai Documenti

Le seguenti sono differenti librerie che possono essere usate per analizzare e manipolare una varietà di formati di documenti.

3.3.1 Xerces

Il progetto Xerces fornisce un parser XML per una varietà di linguaggi, inclusi Java, C++ e Perl. I binding Perl sono basati sui sorgenti C++. Un parser XML è uno strumento usato per l'accesso programmatico a documenti XML. Ecco una descrizione degli standard supportati da Xerces:

- *DOM* `http://www.w3.org/TR/1998/REC-DOM-Level-1-19981001/level-one-core.html` : DOM sta per Document Object Model. I documenti XML sono per natura gerarchici (tag annidati). Ai documenti XML si può accedere tramite un'interfaccia strutturata ad albero. Il processo si svolge così:
  - Analisi del documento
  - Costruzione dell'albero
  - aggiunta/cancellazione/modifica dei nodi
  - Serializzazione della struttura
- *SAX* `http://www.saxproject.org` :Semplice API per XML, basata su flusso; il che significa che lancerà degli eventi in base agli elementi incontrati. Tali eventi possono essere usati per costruire, per esempio, una intelaiatura DOM.
- *XML Namespaces* `http://www.w3.org/TR/REC-xml-names/`
- XML Schema: Lo standard XML fornisce la sintassi per la scrittura di documenti. XML Schema fornisce gli strumenti per la definizione dei *contenuti* del documento XML (semantica). Permette di definire se un certo elemento nel documento debba essere un intero compreso tra 10 e 20 o contenere un indirizzo IP.

L'iniziale codice di base per il progetto XML di Xerces fu donato da IBM. Maggiori informazioni si possono ottenere presso le homepage *Xerces Java* `http://xml.apache.org/xerces-j/index.html` , *Xerces C++* `http://xml.apache.org/xerces-c/index.html` e *Xerces Perl* `http://xml.apache.org/xerces-p/index.html` .

3.3.2 Batik

Batik è un toolkit basato su Java per applicazioni che vogliano usare immagini nel formato

Scalable Vector Graphics (SVG) `http://www.w3.org/TR/SVG/` *per vari scopi come la visualizzazione, la generazione o la manipolazione.*

È incentrato su XML e aderente alle specifiche W3C. È un po' atipico rispetto agli altri progetti Apache, in quanto fornisce un componente grafico. Batik fornisce degli hook per estendere la struttura tramite tag personalizzati e permette la conversione da SVG ad altri formati come JPEG o PNG. Si può saperne di più sulla *homepage di Batik* `http://xml.apache.org/batik/`

3.3.3 POI

Il progetto POI consiste di API per manipolare vari formati di file basati sul formato OLE 2 Compound Document di Microsoft usando soltanto Java. Questo include i documenti Word e Excel. Si possono trovare maggiori informazioni presso

3.4 Interoperabilità

Le seguenti sono librerie per la comunicazione remota e l'interoperabilità tra i server.

3.4.1 SOAP

Apache SOAP (Simple Object Access Protocol) e Axis sono implementazioni del *protocollo SOAP* `http://www.w3.org/TR/SOAP`

SOAP è un protocollo leggero per lo scambio di informazioni in ambiente decentralizzato e distribuito. È un protocollo basato su XML che consiste di tre parti:

- *un involucro che definisce un'ossatura per la descrizione di cosa ci sia in un messaggio e come processarlo*,
- *un insieme di regole di codifica per l'espressione di istanze di tipi di dati definiti dall'applicazione*, e
- *una convenzione per rappresentare chiamate e risposte di procedure remota.*

Fondamentalmente, si può pensare a SOAP come a un sistema di chiamate di procedura remota, basato su HTTP e XML. Da un lato questo comporta verbosità e lentezza rispetto ad altri sistemi. Dall'altro agevola l'interoperabilità, il debugging e lo sviluppo di client e di server per una varietà di linguaggi, dal momento che la maggioranza dei moderni linguaggi dispone di moduli HTTP e XML. Per saperne di più si veda la *homepage di Apache SOAP* `http://xml.apache.org/soap/`

3.4.2 XML-RPC

Il *progetto XML-RPC* `http://xml.apache.org/xmlrpc/`

è un'implementazione Java del protocollo XML-RPC, un protocollo leggero simile a SOAP e suo predecessore.

3.4.3 Sicurezza XML

Il *progetto sicurezza di XML* `http://xml.apache.org/security/` fornisce la verifica della firma digitale dei documenti XML per lo scambio sicuro di documenti.

3.5 Sviluppo

3.5.1 Apache Portable Runtime

Il progetto *APR* `http://apr.apache.org` fornisce uno strato di portabilità che astrae un certo numero di API per la manipolazione di file, l'accesso alla rete e così via. È scritto in C e funziona sulla maggior parte delle versioni di Unix, su Windows e una varietà di altri sistemi. Costituisce la base di Apache 2.0.

3.5.2 Ant

Ant `http://jakarta.apache.org/ant/` è uno strumento di compilazione scritto in Java. Dispone di una API modulare e può essere esteso con la creazione di nuovi task. È governato da file di configurazione in XML.

3.5.3 Byte Code Library

La *Byte Code Engineering Library* `http://jakarta.apache.org/bcel/` (BCEL) è una libreria per analizzare, creare e manipolare i file binari delle classi Java.

3.5.4 Log4j

Questo pacchetto fornisce una infrastruttura per la gestione dei log compatibile con le applicazioni Java. Può essere abilitato all'avvio senza modificare il file binario ed è stato progettato con un occhio alle prestazioni. Può essere trovato presso

3.5.5 ORO and Regexp

ORO è un pacchetto completo che fornisce supporto alle espressioni regolari per Java. Include il supporto alle espressioni regolari Perl5, alle espressioni glob e così via. Il tutto sotto licenza Apache. Per saperne di più su ORO si veda . Un altro pacchetto leggero, della ASF, per espressioni regolari, è *Regexp* `http://jakarta.apache.org/regexp/` .

3.5.6 Struts

Struts è un progetto Apache che tenta di trasportare il paradigma progettuale Model-View-Controller (MVC) nello sviluppo web. È basato sulle tecnologie *Servlet* `http://java.sun.com/products/servlet` e

JavaServer Pages `http://java.sun.com/products/jsp` . La parte modello è composta di oggetti Java lato server, che rappresentano lo stato interno dell'applicazione. La parte vista è costruita attraverso le JavaServer Pages (JSP), che sono una combinazione di HTML/XML statico e di Java. Le JSP permettono inoltre allo sviluppatore di definire nuovi tag. La parte controller consiste di servlet, che accolgono le richieste (GET/POST) del client, agiscono sul modello e aggiornano la vista fornendo la JSP appropriata. È possibile saperne di più sulla *pagina del progetto Struts* `http://jakarta.apache.org/struts/index.html` .

3.5.7 Taglibs

La tecnologia JavaServer Pages permette agli sviluppatori di aggiungere funzionalità tramite tag personalizzati. Il progetto Taglibs intende essere un archivio comune per tali estensioni. Include tag per utilità comuni (per es. date), per l'accesso a database SQL e così via.

Per saperne di più su TagLibs si veda . Una più ampia documentazione è inclusa nel pacchetto.

3.5.8 Database

OJB `http://jakarta.apache.org/ojb/` è uno strumento per la mappatura di database che permette la persistenza e l'archiviazione di oggetti Java nei database relazionali. *Xindice* `http://xml.apache.org/xindice/` è un database nativo XML per l'archiviazione e l'interrogazione di documenti XML.

3.5.9 Commons

Il *progetto Commons* `http://jakarta.apache.org/commons/`

fornisce una grande varietà di componenti Java riutilizzabili con minime dipendenze.

3.6 Testing

I seguenti progetti ASF riguardano il testing e l'analisi delle prestazioni.

3.6.1 httpd-test

Il *progetto httpd-test* `http://httpd.apache.org/test/` fornisce un ambiente di test per il server web Apache e strumenti come *flood* `http://httpd.apache.org/test/flood/` per il test del carico HTTP.

3.6.2 Cactus

Cactus `http://jakarta.apache.org/cactus/` è un ambiente di test per la verifica di codice Java lato server, come Servlet e EJB.

3.6.3 JMeter

È uno strumento per i test scritto in Java, con un'interfaccia grafica. Si può ottenere presso .

3.6.4 Lakta

Lakta `http://jakarta.apache.org/lakta/` è uno strumento end-to-end di testing HTTP.

3.6.5 Watchdog

Il *progetto Watchdog* `http://jakarta.apache.org/watchdog/` è un corredo di set per la convalida delle specifiche Servlet e JavaServer Pages.

4 Dove trovare maggiori informazioni

Risorse aggiuntive relative a Apache

4.1 Siti Web

Alcuni utili siti web

- *Sito Web di Apache* `http://www.apache.org`
- *Apache Week* `http://www.apacheweek.com`
- *directory dei moduli di Apache* `http://modules.apache.org`
- *Apache today* `http://www.apachetoday.com`
- *Apache World* `http://www.apacheworld.org`

- *sezione di Slashdot su Apache* `http://slashdot.org/index.pl?section=apache`

4.2 Libri

L'autore mantiene *una lista di libri* `http://www.apacheworld.org/apache_overview/books/` correlati a questo documento. Non è un elenco esaustivo, ma piuttosto include quei libri che l'autore ha personalmente trovato ben scritti e utili.

4.3 Forum di supporto

È possibile trovare la mailing list degli utenti Apache presso . Esistono liste simili per gli altri progetti documentati qui. Prima di inviare richieste di aiuto si legga il documento delle Frequently Asked Questions. Si può ottenere supporto anche nel newsgroup comp.infosystems.www.servers.unix presso .

Se si desidera un supporto commerciale, si valuti se contattare *Covalent* `http://www.covalent.net` , che fornisce supporto di esperti per Apache (ovviamente, a pagamento). Se si sta usando Apache sotto Linux, il proprio venditore Linux potrebbe avere piani di supporto che comprendano anche Apache.

5 Contattare l'Autore

Potete contattarmi all'indirizzo daniel @ rawbyte.com . Suggerimenti e correzioni saranno graditi, vi prego però di non spedirmi messaggi chiedendomi di risolvere i malfunzionamenti della vostra installazione di Apache. Non ho proprio il tempo di rispondere individualmente a tutti. Se avete bisogno di supporto, fate riferimento alle risorse citate sopra.

5.1 Traduzioni

Chiunque desideri contribuire alla traduzione di questo documento dovrebbe utilizzare il sorgente SGML. Si veda per informazioni. Mandatemi una nota così che io possa essere certo che otteniate la versione più recente.

6 Open Content Open Publication License

Open Publication License Draft v1.0, 8 June 1999 (text version)

6.1 REQUIREMENTS ON BOTH UNMODIFIED AND MODIFIED VERSIONS

The Open Publication works may be reproduced and distributed in whole or in part, in any medium physical or electronic, provided that the terms of this license are adhered to, and that this license or an incorporation of it by reference (with any options elected by the author(s) and/or publisher) is displayed in the reproduction.

Proper form for an incorporation by reference is as follows:

Copyright (c) <year> by <author's name or designee>. This material may be distributed only subject to the terms and conditions set forth in the Open Publication License, vX.Y or later (the latest version is presently available at http://www.opencontent.org/openpub/). The reference must be immediately followed with any options elected by the author(s) and/or publisher of the document (see section VI).

Commercial redistribution of Open Publication-licensed material is permitted.

Any publication in standard (paper) book form shall require the citation of the original publisher and author. The publisher and author's names shall appear on all outer surfaces of the book. On all outer surfaces of the book the original publisher's name shall be as large as the title of the work and cited as possessive with respect to the title.

6.2 COPYRIGHT

The copyright to each Open Publication is owned by its author(s) or designee.

6.3 SCOPE OF LICENSE

The following license terms apply to all Open Publication works, unless otherwise explicitly stated in the document.

Mere aggregation of Open Publication works or a portion of an Open Publication work with other works or programs on the same media shall not cause this license to apply to those other works. The aggregate work shall contain a notice specifying the inclusion of the Open Publication material and appropriate copyright notice.

SEVERABILITY. If any part of this license is found to be unenforceable in any jurisdiction, the remaining portions of the license remain in force.

NO WARRANTY. Open Publication works are licensed and provided as is without warranty of any kind, express or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose or a warranty of non-infringement.

6.4 REQUIREMENTS ON MODIFIED WORKS

All modified versions of documents covered by this license, including translations, anthologies, compilations and partial documents, must meet the following requirements:

- 1. The modified version must be labeled as such.
- 2. The person making the modifications must be identified and the modifications dated.
- 3. Acknowledgement of the original author and publisher if applicable must be retained according to normal academic citation practices.
- 4. The location of the original unmodified document must be identified.
- 5. The original author's (or authors') name(s) may not be used to assert or imply endorsement of the resulting document without the original author's (or authors') permission.

6.5 GOOD-PRACTICE RECOMMENDATIONS

In addition to the requirements of this license, it is requested from and strongly recommended of redistributors that:

- 1. If you are distributing Open Publication works on hardcopy or CD-ROM, you provide email notification to the authors of your intent to redistribute at least thirty days before your manuscript or media freeze, to give the authors time to provide updated documents. This notification should describe modifications, if any, made to the document.

- 2. All substantive modifications (including deletions) be either clearly marked up in the document or else described in an attachment to the document.
- 3. Finally, while it is not mandatory under this license, it is considered good form to offer a free copy of any hardcopy and CD-ROM expression of an Open Publication-licensed work to its author(s).

6.6 LICENSE OPTIONS

The author(s) and/or publisher of an Open Publication-licensed document may elect certain options by appending language to the reference to or copy of the license. These options are considered part of the license instance and must be included with the license (or its incorporation by reference) in derived works.

A. To prohibit distribution of substantively modified versions without the explicit permission of the author(s). Substantive modification is defined as a change to the semantic content of the document, and excludes mere changes in format or typographical corrections.

To accomplish this, add the phrase 'Distribution of substantively modified versions of this document is prohibited without the explicit permission of the copyright holder.' to the license reference or copy.

B. To prohibit any publication of this work or derivative works in whole or in part in standard (paper) book form for commercial purposes is prohibited unless prior permission is obtained from the copyright holder.

To accomplish this, add the phrase 'Distribution of the work or derivative of the work in any standard (paper) book form is prohibited unless prior permission is obtained from the copyright holder.' to the license reference or copy.