

Strumenti Web per il Commercio Elettronico



Moreno Marzolla
Dipartimento di Informatica
Università Ca' Foscari di Venezia
marzolla@dsi.unive.it
<http://www.dsi.unive.it/~marzolla>

L'hardware per il Commercio Elettronico

- Fare commercio elettronico significa principalmente operare sul WEB
- Materiale occorrente:
 - Un collegamento alla rete
 - Un nome di dominio (www.miaditta.it)
 - Una piattaforma:
 - Un server WEB
 - Eventuali altri server e applicazioni specifiche per la gestione di dati, dei pagamenti, degli ordini...

Moreno Marzolla

Tecnologie Web

2

Hosting o non hosting

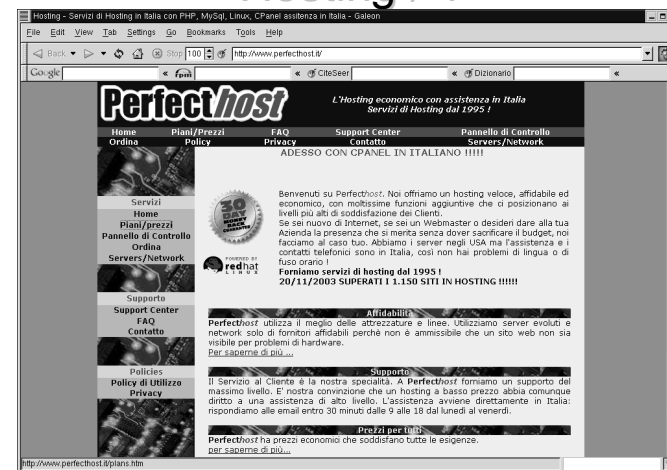
- Domanda: dove ospitare il proprio sito di ECommerce?
 - All'interno della propria azienda?
 - Presso terzi?
- Per rispondere a questa domanda, occorre porsi altre
 - Cosa ci si prefigge?
 - In che misura il personale è coinvolto nella gestione dei vari server?
 - Abbiamo le competenze tecniche per sviluppare una soluzione "in casa"?

Moreno Marzolla

Tecnologie Web

3

Hosting / 1



Moreno Marzolla

Tecnologie Web

4

Hosting / 2

The screenshot shows the Perfecthost website interface. At the top, it says "Perfecthost L'Hosting Perfetto con assistenza in Italia Servizi di Hosting dal 1995!". Below this is a navigation menu with links like Home, Ordina, Servizi, Piani/prezzi, Pannello di Controllo, Ordina, Servers/Network, Supporto, Support Center, Community, FAQ, and Contatto. The main content area features a table titled "Ordina Subito" with columns for "Costo Mensile (IVA esclusa)", "Garanzia 30 giorni soddisfatti o rimborsati", and various service options. The table lists different hosting plans (Entry, Bronzo, Argento, Oro, Platino) and their associated features and costs.

	Entry	Bronzo	Argento	Oro	Platino
Costo Mensile (IVA esclusa)	9,90	12,90	19,90	29,90	
Garanzia 30 giorni soddisfatti o rimborsati	SI	SI	SI	SI	SI
Spazio Disco (MB)	10	100	200	300	500
Registrazione dominio (1)	inclusa	inclusa	inclusa	inclusa	inclusa
Banda disponibile	3 x OC3	3 x OC3	3 x OC3	3 x OC3	3 x OC3
Traffico mensile (MB)	1000	6000	9000	12000	15000
Account FTP	1	3	5	12	30
Database MySQL	-	3	4	5	6
Indirizzi E-mail(*)	2	5	10	15	20
Mailing List	-	1	2	10	20
Indirizzi E-mail(*)	Entry	Bronzo	Argento	Oro	Platino
Apache Web Server	SI	SI	SI	SI	SI
Connessioni OC3 Multiple	SI	SI	SI	SI	SI
Red Hat Linux	SI	SI	SI	SI	SI
Assistenza in Italia	SI	SI	SI	SI	SI
Assistenza in Italia	Entry	Bronzo	Argento	Oro	Platino
Autorisponditori	SI	SI	SI	SI	SI
Alias E-mail	Illimitati	Illimitati	Illimitati	Illimitati	Illimitati
E-mail Forwarding	Illimitati	Illimitati	Illimitati	Illimitati	Illimitati
Gestione Antispam	SI	SI	SI	SI	SI
Filtri sulla posta	SI	SI	SI	SI	SI

Moreno Marzolla

Tecnologie Web

5

Alcune considerazioni su cui riflettere

- Per l'azienda che muove i primi passi nel mondo dell'ECommerce, può risultare conveniente rivolgersi a terzi
 - Mancanza di personale esperto interno
 - Mancanza di conoscenze tecniche specifiche
 - Poco tempo a disposizione per sviluppare soluzioni ad-hoc

Moreno Marzolla

Tecnologie Web

6

A favore del fornitore esterno / 1

- Può essere la soluzione ideale:
 - Le aziende con poca esperienza possono evitare gli alti investimenti derivanti dall'acquisto di hardware e software
 - Si evita di dover creare uno staff di personale qualificato per la gestione del sito
 - Appoggiandosi interamente ad un fornitore esterno, si evita di dover acquistare e gestire una connessione diretta a Internet
 - Gli ISP vendono sia l'accesso a Internet sia le competenze dei propri esperti

Moreno Marzolla

Tecnologie Web

7

La soluzione interna

- Dopo una fase iniziale, la propria attività potrebbe crescere a tal punto che affidare la gestione all'esterno può rivelarsi inefficiente e antieconomico
- ...E' il momento di gestire il sito in proprio!
 - E' più facile e immediato risolvere eventuali problemi (niente interazioni "remote")
 - Le competenze acquisite internamente consentono di eliminare le dipendenze dall'esterno
 - Migliore integrazione con il sistema informativo aziendale

Moreno Marzolla

Tecnologie Web

8

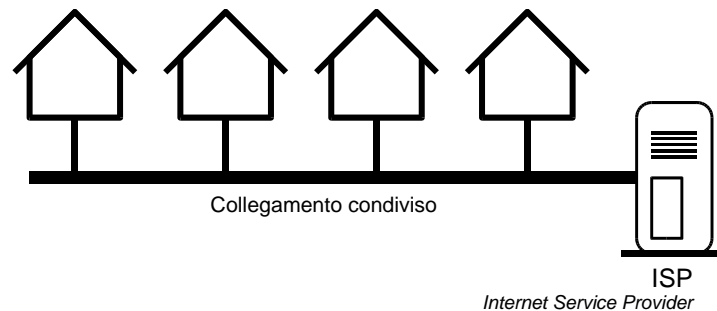
La soluzione interna

- Supponiamo di aver optato per la soluzione interna. Certamente dobbiamo acquistare “connettività” da un ISP
 - Dobbiamo acquistare un accesso alla rete, e un nome di dominio
- Quanto deve essere veloce la nostra connessione alla rete?
 - Sito *molto* frequentato → elevata larghezza di banda richiesta
 - Sito *poco* frequentato → bassa larghezza di banda richiesta

Considerazioni sulla larghezza di banda

- La “larghezza di banda” è la velocità della nostra connessione, misurata in bytes al secondo
 - Inizialmente non ne servirà molta
 - Però il vostro sito potrebbe avere successo. Il vostro ISP è in grado di fornirvi più banda se ne avete bisogno?
 - Il vostro ISP vi *garantisce* la banda che vi mette a disposizione, oppure tale banda è condivisa insieme ad altri utenti?

Banda condivisa



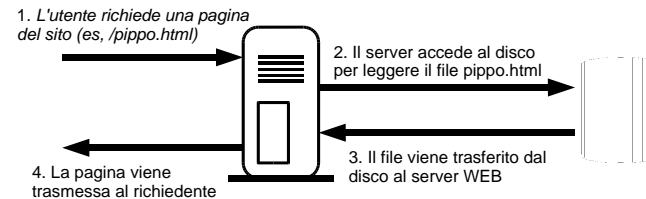
Scegliere la piattaforma

- Un server veloce è preferibile ad uno lento
 - ...però costa anche di più
- Il traffico che ci si aspetta sul proprio sito è un fattore determinante per guidare la scelta della piattaforma
 - *Numero di accessi* all'ora (hit)
 - *Numero di transazioni* (acquisti) completate all'ora
 - *Traffico* di rete (numero di bytes trasmessi/ricevuti all'ora)

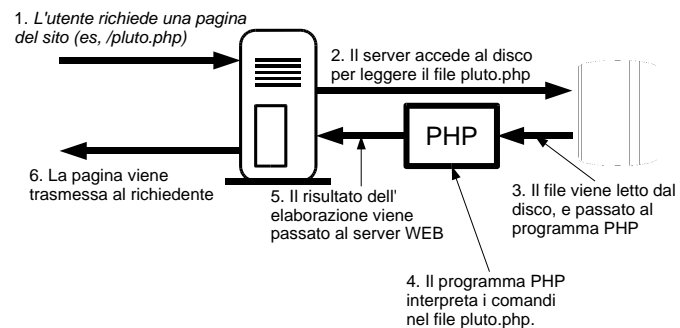
Compiti del server WEB / 1

- Mantiene l'elenco delle pagine disponibili
- Rimane in attesa di richieste da parte degli utenti (client)
- Quando riceve una richiesta:
 - Se è per una pagina statica, la trasmette direttamente al richiedente
 - Se è per una pagina dinamica, la processa e trasmette al richiedente il risultato dell'elaborazione
 - NB: Le pagine dinamiche vengono usate, ad esempio, per leggere il listino dei prodotti direttamente dal database aziendale

Pagine statiche



Pagine dinamiche



Compiti del server WEB / 2

- Devono garantire la sicurezza del sito
 - Autenticare gli utenti tramite password
 - Stabilire connessioni sicure quando richiesto (es. per la trasmissione dei dati del pagamento)
 - Mantenere traccia di tutti gli accessi (data, ora, nome della pagina richiesta, indirizzo IP del richiedente) per una successiva analisi
 - Ad esempio, per individuare quali sono le pagine più visitate dagli utenti

Valutare le prestazioni di un server WEB

- **Benchmarking**
 - Insieme delle attività di prova e di misura effettuate per paragonare le prestazioni di hardware e software.
- E' importante conoscere i fattori che influenzano le prestazioni del server WEB
 - Un server troppo lento scoraggia gli utenti, e causa mancati guadagni

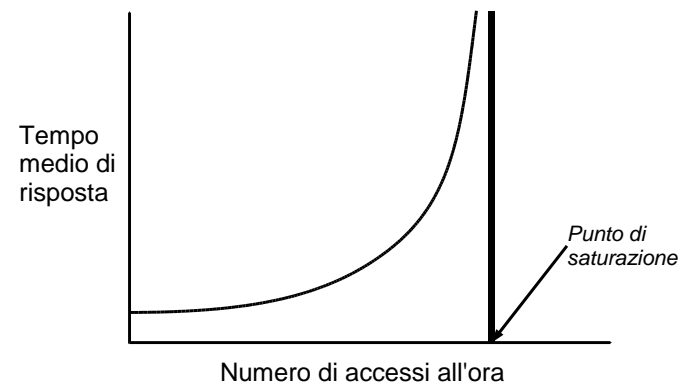
Cosa dobbiamo misurare?

- **Throughput**
 - Numero di richieste a cui si risponde per unità di tempo; più è alto, meglio è
 - Es: "Questa combinazione di hardware/server WEB è in grado di soddisfare 3000 richieste al minuto"
- **Tempo di risposta**
 - Quanto tempo passa tra l'invio di una richiesta e il ricevimento della risposta; più è basso, meglio è
 - Es: "Questa combinazione hardware/server WEB garantisce tempo medio di risposta di 2s"

Fattori che influenzano le prestazioni di un server WEB

- **Caratteristiche dell'hardware**
 - Velocità del processore, quantità di memoria, velocità dei dischi...
- **Velocità della rete**
 - Un server potente collegato tramite un modem fornirà certamente prestazioni scadenti
- **Quantità di utenti che il server deve gestire**
 - Qualunque soluzione, anche la più sofisticata, non può scalare indefinitamente
- **Tipo di richieste**
 - Pagine statiche o dinamiche?

Scalabilità dei server WEB



Strumenti di misura delle prestazioni dei server WEB

- Esistono appositi strumenti (programmi di *benchmarking*) per misurare le prestazioni dei server WEB
- Tali programmi interagiscono con il WEB server, simulando accessi simultanei da parte di più utenti, e registrano i tempi di risposta
- In tal modo è possibile quantificare in modo preciso le prestazioni del server

Prestazioni dei server

- Le prestazioni variano di continuo man mano che macchine più potenti e programmi più aggiornati sono disponibili
- Alcune considerazioni
 - Le prestazioni di alcuni server degradano paurosamente all'aumentare del numero di accessi
 - Vi sono tipi di pagine (pagine dinamiche) che vengono gestite meno efficientemente da parte di alcuni tipi di server
 - Server diversi hanno prestazioni *nettamente* diverse

SPECweb99 / 1

- Un benchmark di tipo client/server per misurare il numero massimo di connessioni simultanee supportate da un web server
- Un insieme di macchine client forniscono un carico di lavoro (workload) al server http

<http://www.spec.org/>

SPECweb99 / 2

- Dalle FAQ su SPECweb99:

"The SPECweb99 workload simulates the accesses to a web service provider, where the server supports the home page for a number of different organizations. Each home page is a collection of files ranging in size from small icons to large documents or images. As in the real world, certain files within the home page are more popular than others. The dynamic GETs simulate the common practice of "rotating" advertisements on a web page. The POSTs simulate entry of user data into a log file on the server, such as might happen during a user registration sequence."

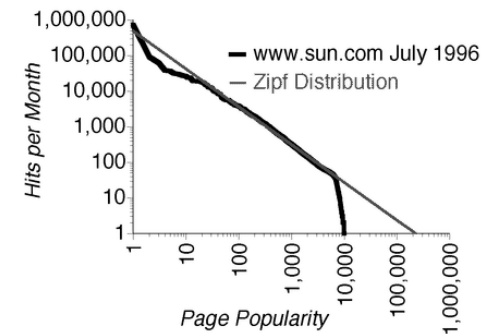
Distribuzioni della popolarità

- La distribuzione degli accessi segue una legge di Zipf:
 - Ogni evento ha una dimensione y e un rango r (rango rappresenta l'ordinamento degli eventi in base alla dimensione)
 - **Legge di Zipf:** La dimensione dell' r -esimo evento è inversamente proporzionale al suo rango: $y \approx r^{-b}$
 - Esempio: y =numero di accessi alle pagine web; ordiniamo le pagine in senso decrescente per numero di accessi; r indica la r -esima pagina più frequente

Esempio

<http://www.useit.com/alertbox/zipf.html>

Popolarità delle pagine del sito www.sun.com



Applicazioni all'eCommerce

- Le pagine più frequentemente visitate:
 - Sono le migliori candidate ad ospitare banner pubblicitari
 - Sono le migliori candidate ad ospitare (rimandi a) informazioni importanti, così l'utente non deve cercarle in giro per il sito
 - Devono essere le più veloci da visualizzare per non sovraccaricare il server WEB

I portali / 1

- Possono essere considerati i “campi base” dai quali gli utenti partono per la propria navigazione o per effettuare ricerche su WEB
 - Forniscono informazioni di tipo generale
 - Non sono mirati a certi tipi di utenti
 - Consentono di effettuare ricerche
 - Spesso possono essere personalizzati per mostrare informazioni di interesse
 - Viaggi, Meteo, Shopping...

I portali / 2

- Forniscono una serie di servizi accessori
 - Caselle E-mail gratuite
 - Notizie di attualità
 - Collegamenti a chat, a negozi virtuali...
 - Servizi aggiuntivi per gli iscritti
- Gli utenti spendono tempo a personalizzare il proprio portale preferito, quindi tendono a rimanervi fedeli
 - La fedeltà degli utenti è la principale fonte di reddito per i portali

Moreno Marzolla

Tecnologie Web

33

Virgilio.it

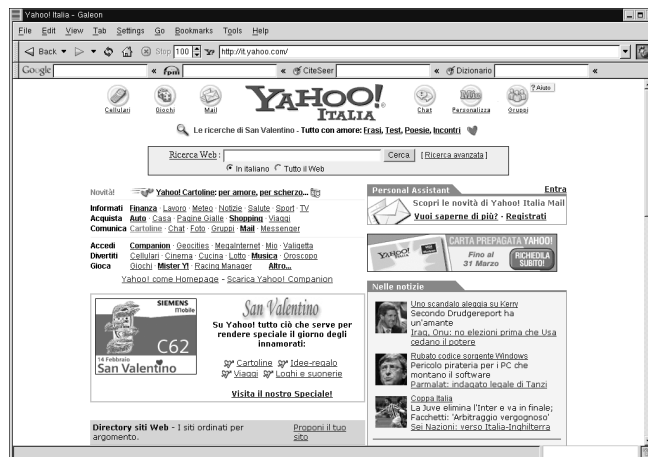


Moreno Marzolla

Tecnologie Web

34

Yahoo.it / 1



Moreno Marzolla

Tecnologie Web

35

Yahoo.it / 2



Moreno Marzolla

Tecnologie Web

36

I motori di ricerca

- E' un particolare tipo di programma che localizza pagine WEB in base al contenuto
 - Es: "Sto cercando tutti i siti che parlano di aeromodellismo"
- Il motore di ricerca non guarda ogni volta tutte le pagine del WEB alla ricerca di quelle che interessano
 - Piuttosto, cerca in un proprio database interno le possibili corrispondenze
 - Il database viene aggiornato periodicamente, visitando tutto (o buona parte) il WEB

Altavista / Google



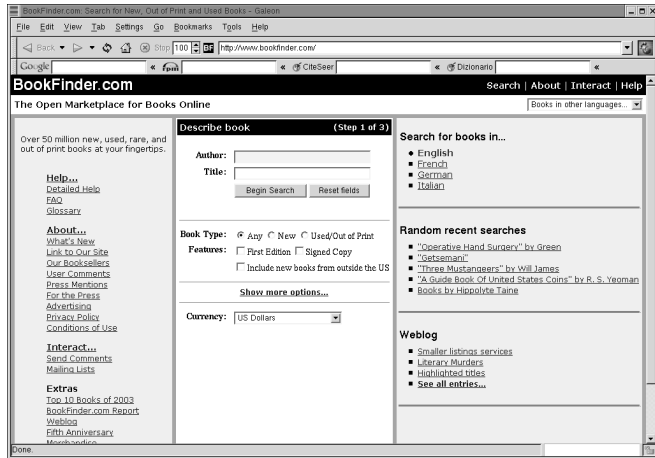
Agenti intelligenti

- L'attuale frontiera del WEB
- Sono programmi che, per conto dell'utente, possono effettuare
 - ricerche di informazioni
 - selezione di informazioni
 - mediazione per conto di una persona o di una entità

Esempi

- *“Voglio acquistare un PC con determinate caratteristiche e prezzo sotto i 1000 euro. Dove lo posso trovare?”*
- *“Avvisami quando il prezzo d'asta di questo prodotto scende sotto i 200 euro”*
- *“Compra per me 100 azioni IBM quando il loro prezzo è sceso del 5% rispetto alla media dell'ultima settimana”*

www.bookfinder.com

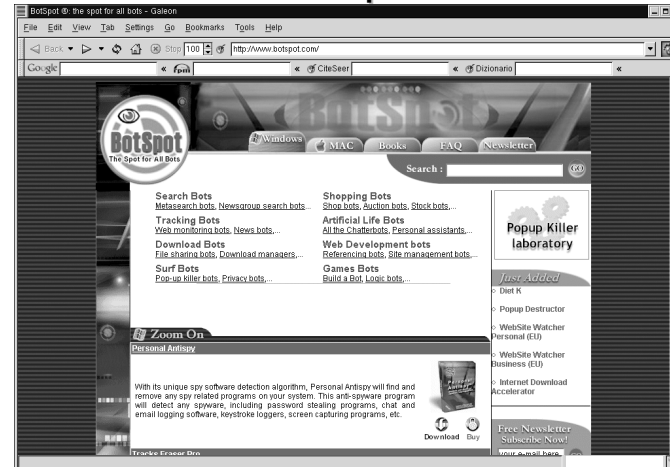


Moreno Marzolla

Tecnologie Web

41

www.botspot.com



Moreno Marzolla

Tecnologie Web

42