

Applicazioni software per il Commercio Elettronico



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

Che tipo di applicazione software è necessaria?

- Le soluzioni software (programmi) e hardware (calcolatori) per la realizzazione di un sito di ECommerce variano molto in base a ciò che ci si prefigge
 - Una grande azienda che vede il WEB come un mercato di sbocco per tutti i suoi prodotti
 - Una piccola realtà che vede il WEB come un mercato accessorio da affiancare alla propria rete di vendita tradizionale
- Soluzioni di costo diverso per obiettivi diversi
 - Da pochi euro al mese a centinaia di migliaia di euro per un prodotto di fascia alta

Come scegliere?

- Dimensione della vostra impresa
 - Quale è il volume stimato delle vendite?
 - Quanti prodotti volete inserire nel catalogo elettronico?
 - Quanto volete spendere per l'attività di ECommerce?
- Una buona notizia
 - Qualunque sia la vostra scelta, molto probabilmente iniziare una attività di ECommerce costerà *meno* dell'apertura di un punto vendita tradizionale
 - Punto vendita tradizionale: Affittare un locale, spese di arredamento, spese di manutenzione
 - Sito ECommerce: affidarsi ad un ISP esterno per tutto (oppure acquistare macchine, software e connettività alla rete)

Ancora sulla scelta del software

- Il software da utilizzare per una attività di ECommerce *business-to-business* (B2B) è diverso da quello che si usa nelle attività *business-to-customer* (B2C)
 - B2B: Attività commerciale tra commercianti (grossista-commercianti, produttore-distributori...)
 - B2C: Attività commerciale diretta al consumatore

Requisiti per ECommerce

- Presentazione del catalogo
- Gestione del carrello elettronico
- Gestione della transazione (pagamento)
- Strumenti per la gestione del catalogo
 - Aggiornamento elenco prodotti e prezzi, offerte speciali...

Moreno Marzolla

Tecnologie Web

5

La presentazione del catalogo

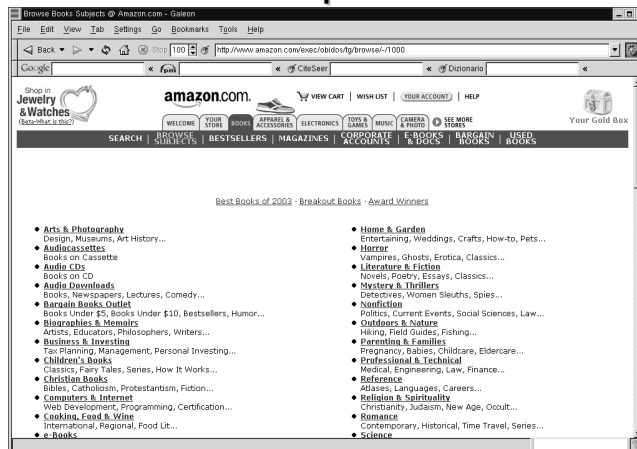
- Aziende con pochi prodotti possono accontentarsi di un catalogo molto semplice
 - Basta una semplice lista dei prodotti
 - Può andar bene per un max di 35 prodotti
- Cataloghi più ampi sono di solito più complessi
 - Foto del prodotto
 - Descrizione dettagliata
 - Prezzo sempre bene in vista
 - Un modo per effettuare ricerche di prodotti
 - Un sistema di navigazione sofisticato

Moreno Marzolla

Tecnologie Web

6

Esempio / 1

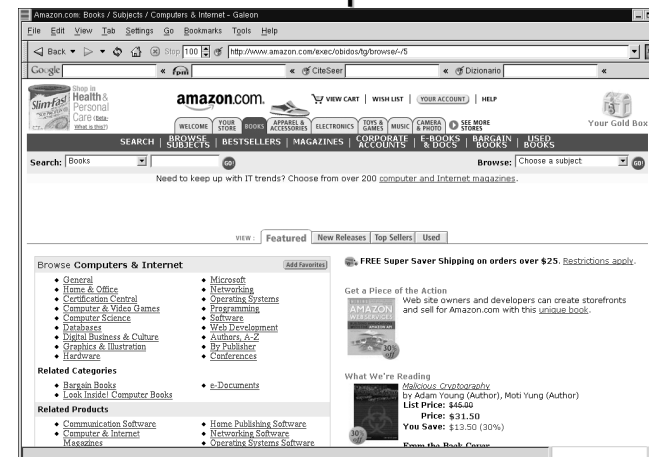


Moreno Marzolla

Tecnologie Web

7

Esempio / 2

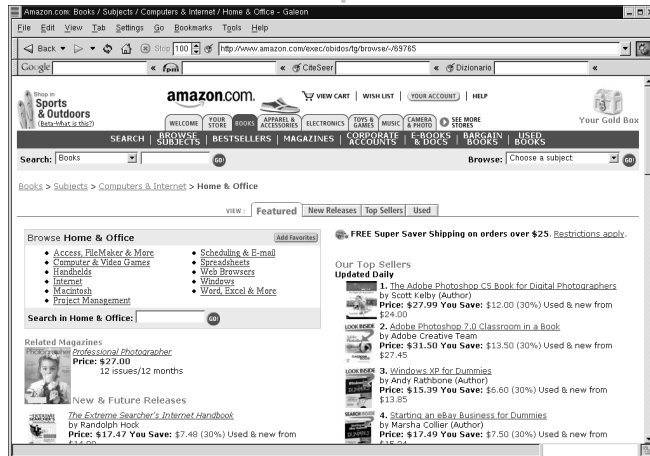


Moreno Marzolla

Tecnologie Web

8

Esempio / 3



Moreno Marzolla

Tecnologie Web

9

Fornire modi diversi per individuare un prodotto

- Ricerca tramite parola chiave nel catalogo
 - Es: "Camicie da uomo"
- Visualizzazione di categorie di prodotti, come nei negozi tradizionali si suddividono i prodotti in diversi *reparti*
 - Es: Abbigliamento sportivo, abbigliamento uomo/ donna/ ambino, abbigliamento per il tempo libero...
- Regola d'oro:

Mai intralciare la strada al cliente che vuole acquistare qualcosa

Moreno Marzolla

Tecnologie Web

10

Effettuare ordini

- Inizialmente gli ordini si facevano compilando una scheda (form) indicando
 - I prodotti che si vogliono acquistare
 - I dati personali del compratore
 - I dati relativi al pagamento
- L'ordine elettronico somiglia molto ai tradizionali ordini cartacei, con gli stessi limiti
 - L'utente deve trascrivere i codici dei prodotti e i prezzi unitari
 - E' molto facile commettere ogni genere d'errore

Moreno Marzolla

Tecnologie Web

11

Esempio

Shipping Information

Ship the order to me
 Ship to the address below

Name:
 Address:
 City, State, Zip:

Order Information: (Please complete entire row)

Coffee Name	Type	Grind	Qty.	Price
Zimbabwe	Caffeinated	Whole Bean	2 lbs	
Yemen Mocha	Caffeinated	Whole Bean	4 lbs	
Kopi Luwak	Caffeinated	Perc	1/2 lb	

S&H: 4.95
 Total:

Comments

NOTE: Please do not enter your credit card number. We will call back to verify the order and receive the credit card number.

Credit Card Information

Name on credit card	Card Type	Expires
<input type="text"/>	<input type="text"/>	<input type="text"/>

Caselle di testo da riempire

Fonte: Schneider e Perry, **Commercio elettronico**, Apogeo

Moreno Marzolla

Tecnologie Web

12

Il carrello elettronico

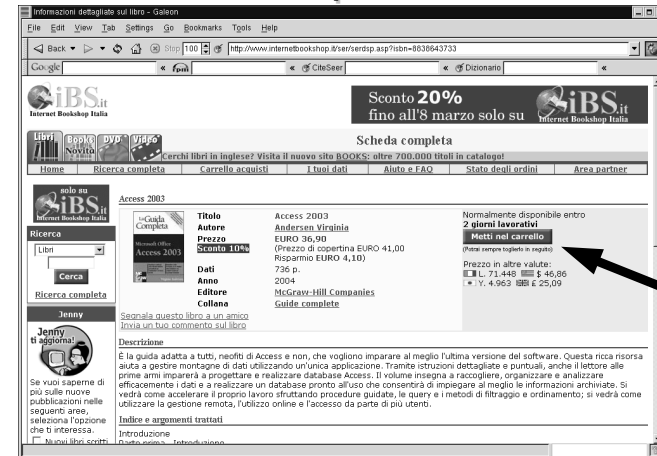
- E' diventato il meccanismo standard con cui si selezionano i prodotti
 - Il cliente "clicca" sul prodotto che vuole acquistare
 - Il carrello elettronico si ricorda la lista dei prodotti selezionati
 - E' possibile rimuovere prodotti dal carrello in qualsiasi momento
 - In certi casi è possibile memorizzare il carrello degli acquisti, per reconsiderarlo in un secondo momento
 - Quando il cliente ha finito, conferma con un click l'acquisto di tutti i prodotti

Moreno Marzolla

Tecnologie Web

13

Esempio / 1

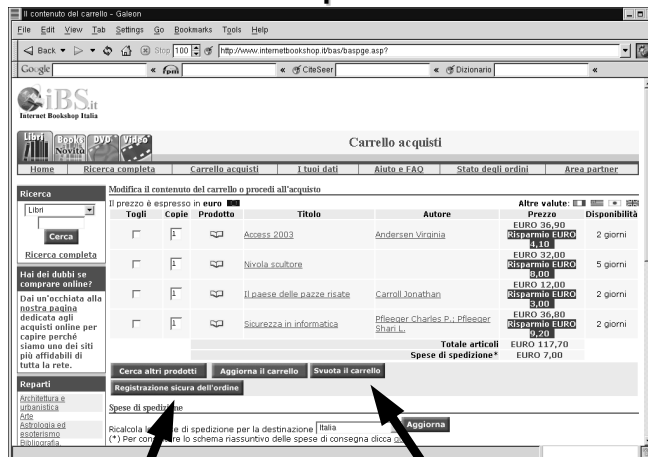


Moreno Marzolla

Tecnologie Web

14

Esempio / 2



Moreno Marzolla

Tecnologie Web

15

La transazione / 1

- La gestione della transazione avviene quando il cliente conferma il contenuto del carrello e procede al pagamento della merce
 - A questo punto il browser entra in modalità "sicura" per garantire che nessuno possa osservare i dati in transito
 - ...Leggi, il numero della nostra carta di credito
 - ...e anche per evitare che i dati dell'ordine vengano alterati

Moreno Marzolla

Tecnologie Web

16

La transazione / 2

- L'elaborazione della transazione è la parte più complicata dell'acquisto
 - E' necessario calcolare il costo dell'ordine, tenendo conto di IVA, sconti, offerte speciali ("sconto del 10% per acquisti superiori ai 100 Euro"), spese di trasporto...
 - Bisogna prestare attenzione perché alcune vendite tra grossisti (Business-to-business) non sono soggetti a imposte
 - Senza contare il commercio con acquirenti esteri...
 - Molti spedizionieri (UPS, FedEx...) consentono ai commercianti un collegamento diretto al proprio sito per calcolare le spese di trasporto
 - Ricordate le Extranet? E il commercio B2B?

Gli strumento per il commercio elettronico

- Alcuni siti (Yahoo) mettono a disposizione soluzioni "chiavi in mano"
 - Creazione cataloghi, gestione carrello elettronico, hosting...
- In altri casi, è necessario dotarsi di software e hardware adeguati, o svilupparsi una qualche soluzione "in casa"

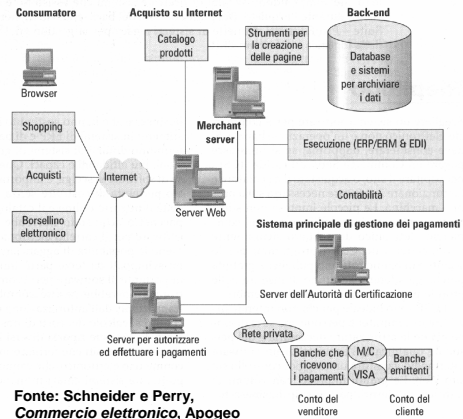
Soluzioni di base

- Alcuni siti di ECommerce offrono soluzioni gratuite o a basso costo per le aziende di piccole dimensione
 - che vendono pochi prodotti (circa meno di 50)
 - che effettuano poche transazioni (basso numero di ordini effettuati)
- Servizi in questa categoria sono di solito basati su form (niente carrello elettronico)
- Spesso non consentono gestione dei pagamenti
- Questi siti guadagnano tramite banner

Soluzioni professionali

- Sono generalmente più costose, ma offrono molti più servizi
- Offrono possibilità di interfacciamento con sistemi informativi aziendali già esistenti
 - Interfacciamento con programmi di gestione basi di dati, magazzino ecc.
 - Interfacciamento con i sistemi informativi aziendali dei fornitori, per soluzioni Business-to-Business
- Offrono funzionalità sofisticate
 - Ad esempio, il riordino automatico dei prodotti tramite Internet quando, a seguito degli acquisti, la giacenza scende sotto un certo livello

Esempio di sistema ECommerce di tipo professionale



Fonte: Schneider e Perry, *Commercio elettronico*, Apogeo

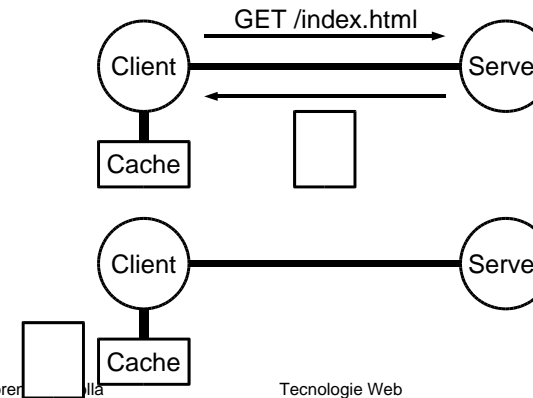
Moreno Marzolla

Tecnologie Web

21

Il protocollo HTTP

- Schema generale di una richiesta



Moreno Marzolla

Tecnologie Web

22

Struttura di una richiesta

```
<METHOD> <URL> <HTTPVERSION>\r\n
<HEADERNAME>: <HEADERVERVAL>\r\n
<HEADERNAME>: <HEADERVERVAL>\r\n
...
\r\n
<DATA, IF POST>
```

```
GET / HTTP/1.1
Host: www.amazon.com
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.7.3) Gecko/20041007 Galeon/1.3.18 (Debian package 1.3.18-1.1)
Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,image/png,*/*;q=0.5
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*/*;q=0.7
Cookie: skin=; session-id-time=1109232000; session-id=102-0987903-6592124; obidos_path_continue-shopping=continue-shopping-url=/subst/home/home.html/102-0987903-6592124&continue-shopping-post-data=&continue-shopping-description=generic.gateway.default; ubid-main=077-6286623-0470965
Connection: close
```

Moreno Marzolla

Tecnologie Web

23

Struttura di una risposta

```
<HTTPVERSION> <STATUS CODE> <MSG>\r\n
<HEADERNAME>: <HEADERVERVAL>\r\n
<HEADERNAME>: <HEADERVERVAL>\r\n
...
\r\n
<DATA, IF NECESSARY>
```

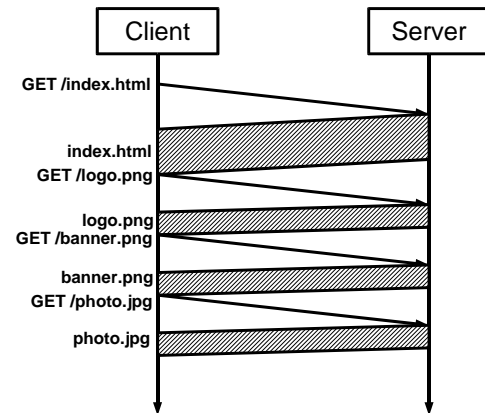
```
HTTP/1.1 301 Moved Permanently
Date: Thu, 17 Feb 2005 09:09:50 GMT
Server: Stronghold/2.4.2 Apache/1.3.6 C2NetEU/2412 (Unix) amarewrite/0.1 mod_fastcgi/2.2.12
Set-Cookie: skin=; domain=amazon.com; path=/;
Location: http://www.amazon.com:80/exec/obidos/subst/home/home.html
Transfer-Encoding: chunked
Content-Type: text/plain
Connection: close
```

Moreno Marzolla

Tecnologie Web

24

Schema HTTP/1.0



Connessioni persistenti

- **Persistent HTTP**
 - in HTTP/1.0, con header "Connection: Keep-Alive\r\n"
 - in HTTP/1.1, P-HTTP built in
- **Pro**
 - consente di inviare più di una richiesta nella stessa connessione TCP (niente penalità per inizializzare nuove connessioni ad ogni richiesta)
- **Contro**
 - Come si fa a sapere quando la pagina è stata ricevuta?
 - Quand'è che si chiude la connessione TCP?
 - Timeout lato server
 - Chiusura della connessione lato client

HTTP/1.0 Metodi

- **GET**
 - Richiede un documento
- **POST**
 - invia i dati di un form
- **HEAD**
 - Simile a GET, ma ritorna solo l'header HTTP e non i dati richiesti. Utile per verificare la presenza di una pagina, o per la gestione della cache lato client
- **PUT, DELETE, LINK, UNLINK**
 - Non usati, grossi problemi di sicurezza

HTTP/1.0 Headers / 1

- **Authorization**
 - Inviato dal client
 - **Authorization: <credentials>**
 - "Basic Auth" is commonly used
 - <credentials> = Base64 (username:password)
 - Ha senso se all'interno di una connessione SSL
- **Content-Encoding**
 - Inviato dal client o dal server
 - **Content-Encoding: x-gzip**

HTTP/1.0 Headers / 2

- **Content-Length**
 - Inviato dal client o dal server
 - **Content-Length: 56**
 - Indica quante informazioni vengono trasmesse
 - Indispensabili per connessioni HTTP persistenti, o quando si usa POST
- **Content-Type**
 - Inviato dal server
 - **Content-Type: text/html**
 - what MIME type the payload is

HTTP/1.0 Headers / 3

- **Date**
 - **Date: Tue, 15 Nov 1994 08:12:31 GMT**
- **Expires**
 - Inviato dal server
 - **Expires: Thu, 01 Dec 1994 16:00:00 GMT**
 - Indica quando "scade" l'informazione trasmessa (utile per le cache)
- **If-Modified-Since**
 - Inviato dal client
 - **If-Modified-Since: Sat, 29 Oct 1994 19:43:31 GMT**
 - Il server restituisce i dati se sono stati modificati dalla data indicata, altrimenti "304 Not Modified"

HTTP/1.0 Headers / 3

- **Last-Modified**
 - Restituito dal server
 - **Last-Modified: Sat, 29 Oct 1994 19:43:31 GMT**
 - Data dell'ultima modifica della risorsa richiesta
- **Location**
 - Restituito dal server
 - **Location: http://www.dsi.unive.it**
 - Usato per le redirezioni
- **Pragma**
 - Inviato dal client o dal server
 - **Pragma: no-cache**

HTTP/1.0 Headers / 4

- **Referer**
 - Inviato dal client
 - **Referer: http://www.xxx-smut.com**
 - Specifica l'indirizzo di provenienza. **Occhio alla privacy!**
- **Server**
 - Restituito dal server
 - **Server: Apache-2.4**
- **User-Agent**
 - Inviato dal client
 - **User-Agent: Mozilla/4.07 [en] (X11; I; Linux 2.0.36 i686)**
 - Perché? Ottimizzare il layout in base alle capacità del client

HTTP/1.0 Headers / 5

- **WWW-Authenticate**
 - Inviato dal server
 - **WWW-Authenticate: <challenge>**
 - Indica al client di reinviare la richiesta con l'header "Authorization:"

Cookies

- Informazioni inviate dai server web
 - Tipicamente 4KB
- Vengono mantenute sull'hard disk dell'utente, e possono essere letti solo dal sito che li ha inviati
- Utilizzati per
 - Gestione delle sessioni;
 - Tracking nei siti
 - Ordini online
 - Pubblicità "su misura"
- Possono essere disabilitati
- <http://www.cookiecentral.com/>

Come funzionano in pratica?

- Il server invia gli header seguenti:

```
...  
Content-type: text/html  
Set-Cookie: foo=bar; path=/; expires Mon, 09-Dec-2002 13:46:00 GMT
```

- Setta un cookie con le seguenti caratteristiche:
 - Il cookie si chiama "foo"
 - Il suo contenuto e' "bar"
 - Il cookie e' valido per l'intero sito (path=/) (alle pagine al di fuori del path *non* viene inviato il cookie)
 - Scade il 9 dicembre 2002
 - Opzionalmente, si puo' settare il dominio di validità del cookie (a quali siti web viene spedito il cookie)

Come funzionano in pratica?

- Il browser, quando accede ad un sito (e ad un percorso) per cui è settato un cookie, invia il seguente header:

```
...  
Cookie: foo=bar  
...
```

- Questo informa il server della presenza del cookie chiamato "foo" di valore "bar"

Cookies

1. Il browser richiede una pagina web
2. Il server invia la pagina e il comando di memorizzare un cookie nel client
3. Il browser memorizza il cookie
4. Ogni volta che viene richiesta una pagina, il browser controlla se esiste un cookie per quel server e per quella pagina
5. In caso affermativo, il browser trasmette il valore del cookie assieme alla richiesta per la pagina
6. Il server riceve la richiesta con associato il cookie, e può utilizzare tale informazione per "ricordarsi" qualcosa circa l'utente

Esempio: usare i cookies per la pubblicità "su misura"

- Visito l'ipotetico sito *www.animalidomestici.it* e seleziono la pagina "Criceti"
- Ricevo la pagina richiesta, che contiene anche un ipotetico banner:

```

```

- Il sito *www.banner.it* invia l'immagine e un cookie, ad esempio:

```
UserPref=Criceti
```

Esemipo / 2

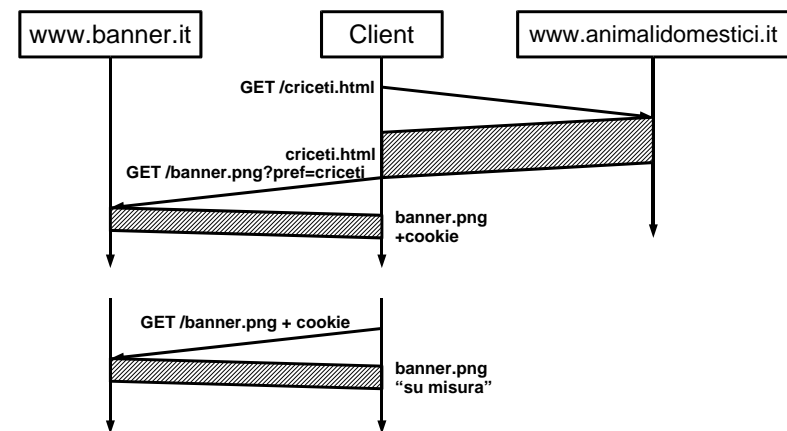
- Da questo momento, ogni volta che dal mio browser ricevo la risorsa generica

```

```

il sito *www.banner.it* mi invierà immagini che hanno a che fare con i criceti

Esempio / 3



HTTP/1.1

- Diversi problemi associati con HTTP/1.0
 - Consistenza delle cache gestite in modo scadente
 - Difficile implementare multi-homed servers
 - 1 indirizzo IP che corrisponde a diversi siti web
 - Difficile precalcolare il content-length
 - Contenuto dinamico...
 - Se la connessione cade, i dati sono perduti
- HTTP/1.1 risolve in parte questi problemi
 - ...a costo dell'introduzione di un notevole livello di complessità
 - HTTP/1.1 NON è un protocollo semplice da implementare

Che c'è di nuovo in HTTP/1.1

- **Host: `www.dsi.unive.it`**
 - Header obbligatorio per il client.
 - Risolve il problema dei siti multi-homed
- **Range: `bytes=300-304,601-993`**
 - Utile se le connessioni si interrompono
- **Age: `<seconds, date>`**
 - Somme dei tempi di vita nelle cache + tempo di trasporto
- **Etag: `fa898a3e3`**
 - Tag univoco per identificare il documento
- **Cache-control: `<command>`**
 - Utile per evitare che la risorsa venga messa nelle cache
- **"chunked" transfer encoding**
 - Documenti "segmentati". Utile nel caso di pagine generate dinamicamente.

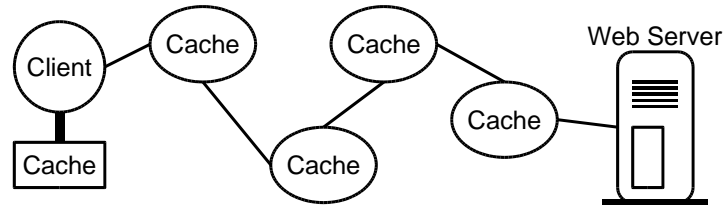
Codici di risposta

- **2xx**
 - Successful
- **3xx**
 - Redirected
 - 301 Moved Permanently
 - 307 Temporary Redirect
- **4xx**
 - Client error
 - Bad request, payment required, forbidden, unauthorized, not found...
- **5xx**
 - Server error

Gestione delle cache

- Problema iniziale del web: nessuna gestione della località
- Risultati: server web che forniscono milioni di volte la stessa pagina a comunità localizzate di utenti
 - Es: gli studenti di una università che consultano le pagine web dei docenti, che magari sono ospitate su un server esterno all'università

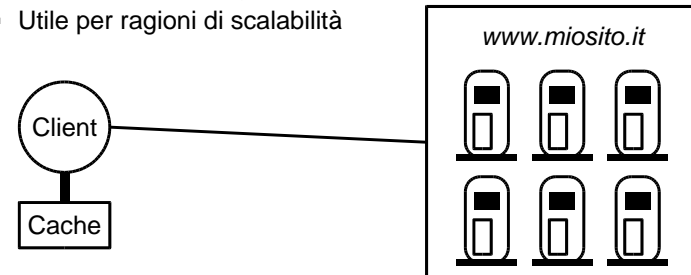
Soluzione: gerarchie di cache



- Si veda www.squid-cache.org

Supporto per server multipli

- Abbiamo visto come gestire diversi server web associati ad un singolo indirizzo IP
 - Utilizzando l'header "Host:" di HTTP/1.1
- Che succede se abbiamo diversi server web che forniscono pagine dello stesso sito web?
 - Utile per ragioni di scalabilità

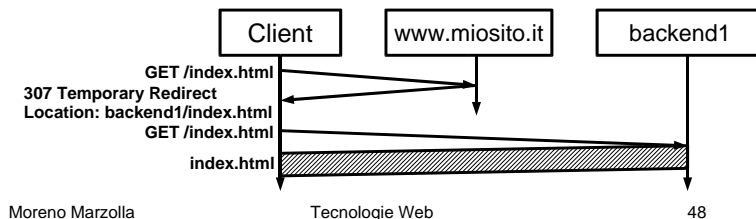


Soluzioni / 1

- DNS round-robin
 - Si assegnano più indirizzi IP ad un singolo nome di dominio
 - Lo stesso nome viene risolto nei diversi indirizzi IP a rotazione
 - Limitazioni
 - I singoli nodi sono esposti direttamente ai client
 - Non tiene conto di fattori come il carico o la potenza di calcolo dei singoli nodi

Soluzioni / 2

- Front-end redirection
 - Un singolo front-end utilizza gli header per la redirection ("temporary redirect") per dirottare le richieste verso i nodi di backend
 - Limitazioni
 - Il front-end è un punto singolo di fallimento
 - Introdotti dei ritardi per la redirection (effetto ping-pong)



Soluzioni / 3

- IP-level multiplexing through smart router
 - Opera a livello IP per indirizzare i pacchetti a host diversi
 - Cisco, SUN, etc. producono hardware specifico per fare ciò
 - Cisco LocalDirector
 - Problematica la gestione dei fallimenti e dello stato
- Soluzioni basate sui client
 - Es, codice Javascript seleziona il proxy più vicino
 - Non basato su HTTP, non tutti i client supportano tali meccanismi

Altre cose da considerare

- Il web non è solo per le persone
 - spiders, crawlers, worms, etc, possono dare problemi
 - Burst di richieste, cicli infiniti nelle visite
- Netscape, IE, ed Apache sono gli standard de facto
 - La loro semantica può differire dagli standard
 - Occorre adeguarsi

Aspetti positivi del WWW

- La gente lo usa per pubblicare informazioni
 - Esistono gli standard
 - Facile “pubblicare” ipertesti
 - Economico
 - Potenzia le facoltà di pubblicazione individuale
- Accessibile
 - Facile da usare/economico
 - Grafica+testo
- Favorisce la condivisione di informazioni
 - Biologia, fisica...

Aspetti positivi del WWW

- Si possono effettuare ricerche
 - Per parola chiave
 - Grande bacino d'utenza, la struttura dei link o i pattern di accesso possono guidare le metriche di rilevanza
 - E' possibile (ma non facile) effettuare query “imprecise”
- Meccanismo uniforme per individuare le risorse
 - URLs
 - Decentralizzato
- Formato standard HTML/CSS
- Il software per l'accesso al WWW è fornito di default

Aspetti positivi del WWW

- Services / Commerce / Information
 - Shopping, viaggi, compravendita di azioni...
 - Aggiornato 24x7
 - Quando e dove vi pare
 - Si possono acquistare prodotti a prezzi convenienti
- Costruisce comunità distribuite
- Le informazioni possono essere presentate in maniera personalizzata

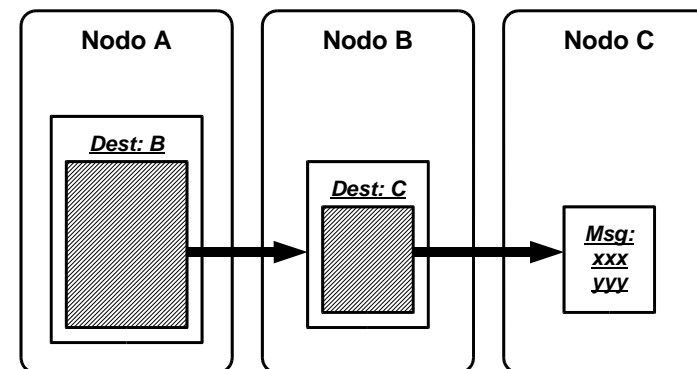
Aspetti negativi del WWW

- Prestazioni non prevedibili
 - "world wide wait"
- Nessun controllo della qualità
 - Tantissime informazioni, alta varianza in qualità e utilità delle informazioni pubblicate
- Mancanza di classificazioni
 - Informazioni pubblicate più velocemente di quanto siano classificate
- La sicurezza è stata aggiunta a posteriori
 - Non è mai una cosa positiva
 - Informazioni personali, anche sensibili, trasmesse elettronicamente

Aspetti negativi del WWW

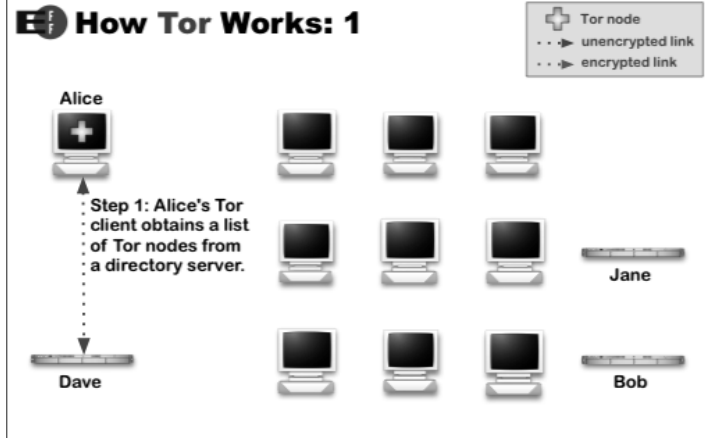
- Totale assenza di privacy
 - Nell'implementazione "di default"
 - La privacy può essere implementata "in cima" al web
 - Onion Routing (www.onion-router.net)
 - Tor (tor.eff.org)
- Onion Routing: idea di base
 - Ciascun nodo possiede una sua chiave crittografica
 - Ciascun nodo è in grado di decifrare l'id del prossimo nodo della catena...
 - ...ma non è in grado di decifrare l'id dei successivi, né può sapere dove è transitato il messaggio prima di arrivare a lui
 - Anche in TV: "Portami dal tuo capo", episodio della serie "Agente Speciale"

Esempio



Tor / 1

How Tor Works: 1



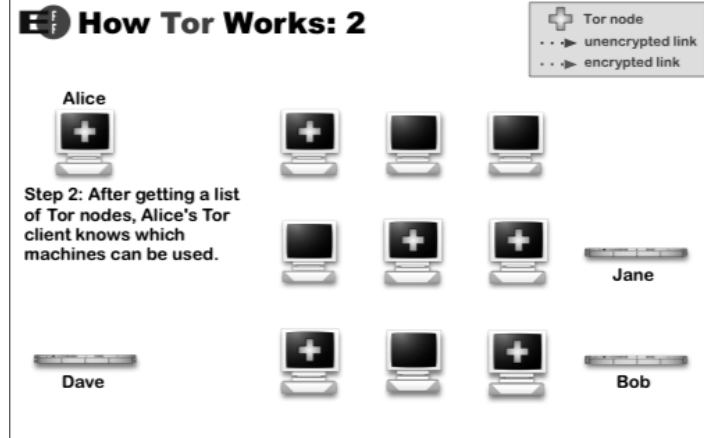
Moreno Marzolla

Tecnologie Web

57

Tor / 2

How Tor Works: 2



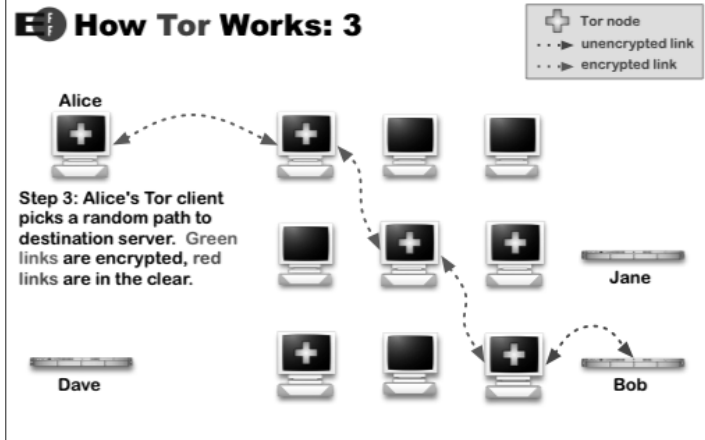
Moreno Marzolla

Tecnologie Web

58

Tor / 3

How Tor Works: 3



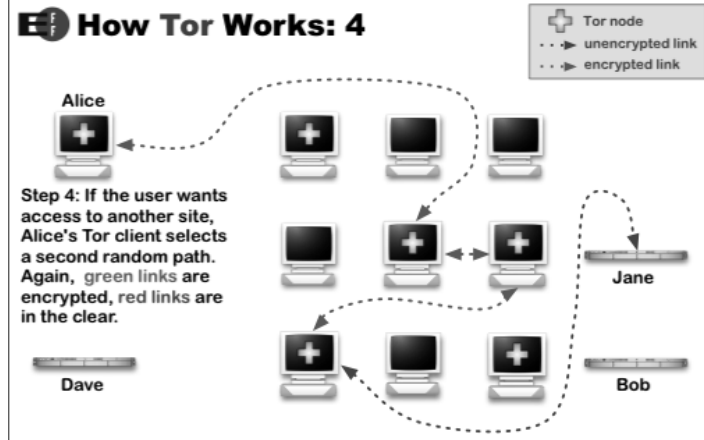
Moreno Marzolla

Tecnologie Web

59

Tor / 4

How Tor Works: 4



Moreno Marzolla

Tecnologie Web

60

Aspetti negativi del WWW

- La gente abusa dell'HTML
 - Gli standard non vengono rispettati, ma le pagine *devono* potersi visualizzare lo stesso
- "eyeball encoding" delle informazioni
 - Le informazioni vengono codificate per l'uso da parte delle persone, non delle macchine
- Dati malamente/non aggiornati
 - Link interrotti

Aspetti negativi del WWW

- Difficile trovare ciò che si vuole
 - Diversi modi per avere risultati simili
 - Difficile specificare le query "giuste"
 - Difficile sapere quanto l'informazione è aggiornata
 - Difficile valutare la completezza dei risultati delle ricerche
 - "Trappole" per aumentare la popolarità
 - Difficile/impossibile ricercare risorse non testuali
 - Alcune ricerche di immagini, es Google

Aspetti negativi del WWW

- Transazioni lunghe sono la norma
 - Nessuna garanzia di consistenza delle transazioni
 - Nessuna conferma end-to-end
- Fattori sociopolitici
- Replication / Caching / Multicast sono problemi ancora presenti e non completamente risolti
- Traffico altamente variabile