

Sistemi Web distribuiti geograficamente e sistemi per Web content delivery

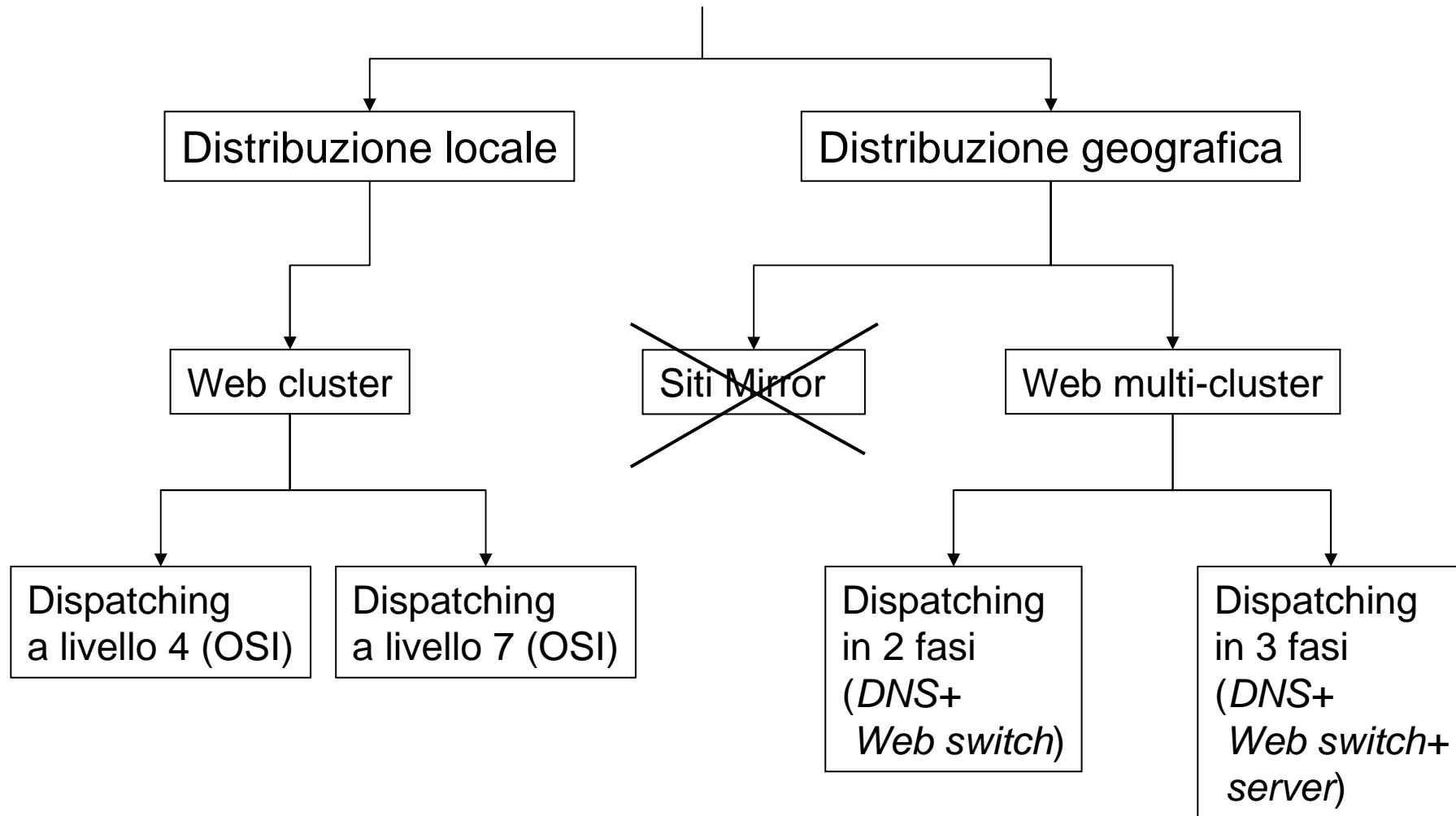
Valeria Cardellini

Università di Roma Tor Vergata

Delivery su scala geografica

- Il content/service provider ha due possibilità per distribuire i propri contenuti/servizi su scala geografica in modo efficiente e scalabile:
 - Il provider possiede e gestisce l'intera piattaforma (**Sistemi Web distribuiti geograficamente**)
 - Il provider gestisce solo i contenuti/servizi ma delega ad una terza parte il servizio di delivery agli utenti (**Content Delivery Network**)

Sistemi con server Web multipli



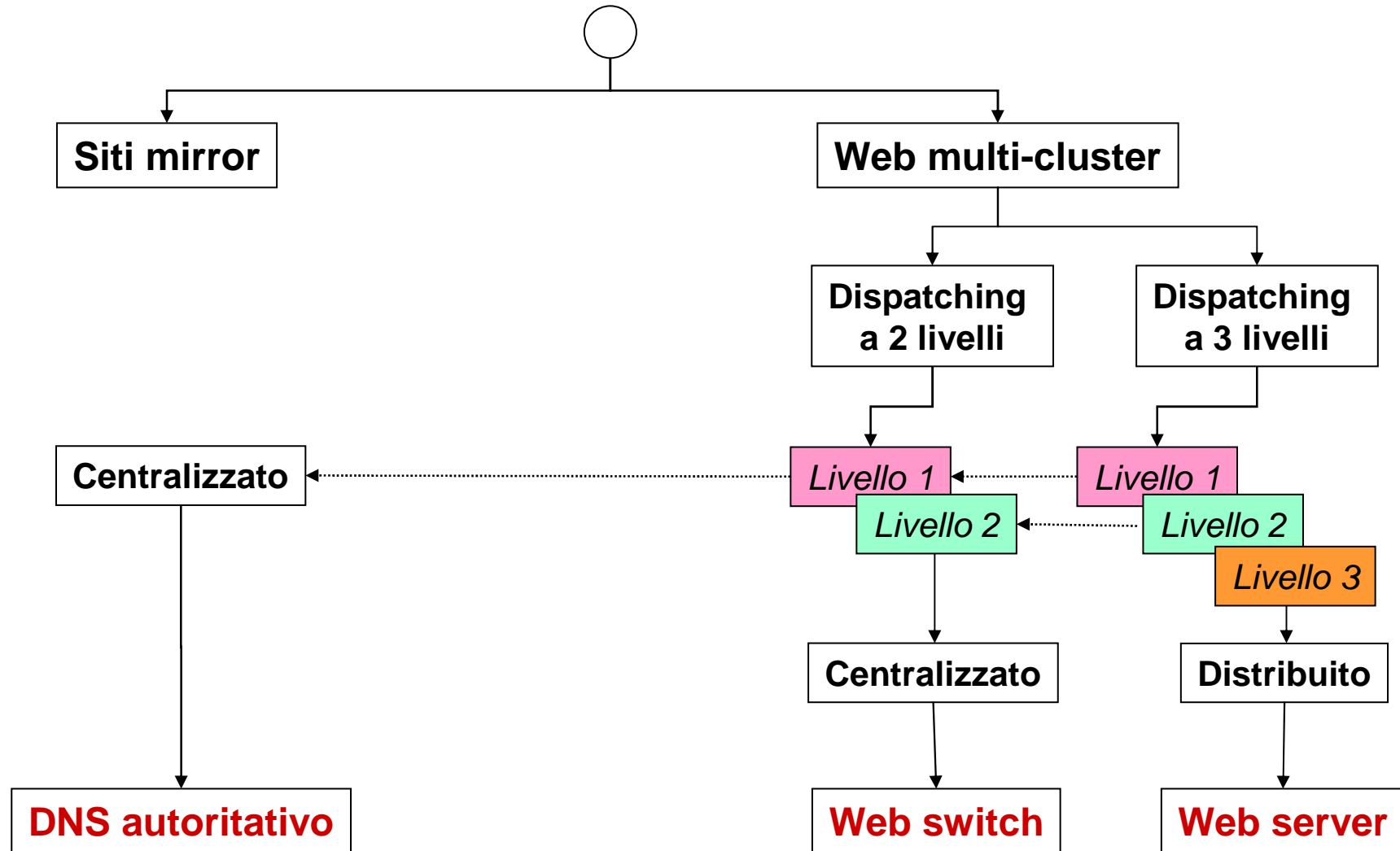
Sistemi Web distribuiti geograficamente

- Problemi dei sistemi Web distribuiti localmente
 - Scalabilità del sistema limitata dalla banda di accesso ad Internet (es., T3 \approx 45 Mbps)
 - Incapacità di evitare i link di rete congestionati
 - Affidabilità della rete
- 
- Scale-out globale
 - Maggiore complessità dell'architettura
 - Meccanismi di routing ed algoritmi di distribuzione
 - Difficoltà nella gestione dell'infrastruttura
 - Metrica per la selezione del cluster “migliore”
 - Localizzazione dei cluster

Web multi-cluster

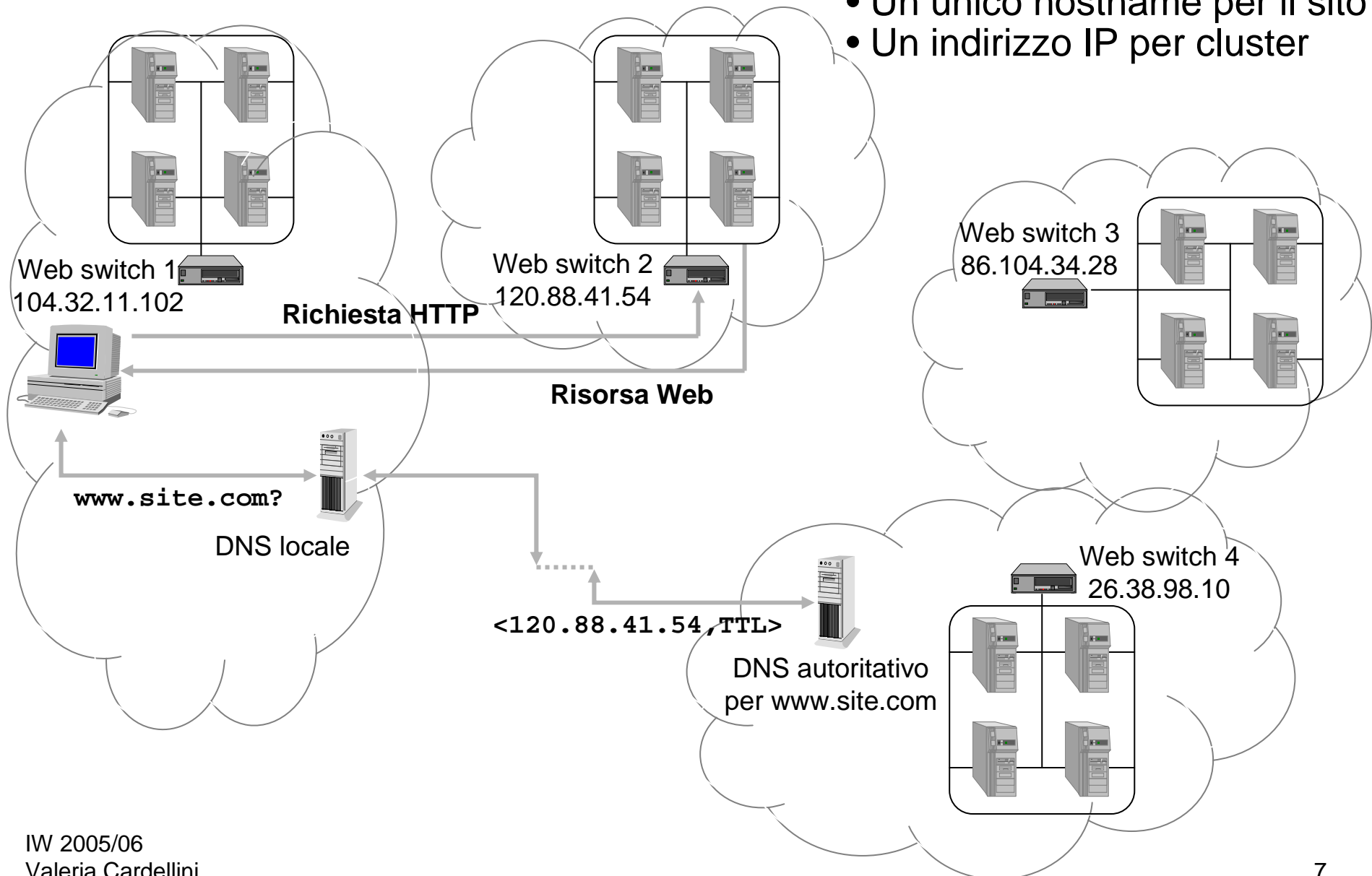
- Sito Web implementato su di un'architettura di Web cluster distribuiti geograficamente tra diverse regioni Internet
- Indirizzi del sito Web
 - Un solo hostname (es., “www.uniroma2.it”)
 - Un indirizzo IP (*VIP dello switch*) per ogni Web cluster
- Il DNS autoritativo del sito Web seleziona un cluster nella *fase di address lookup* mediante un algoritmo di tipo:
 - Round-robin
 - Prossimità geografica
 - Carico dei cluster
 - Prossimità geografica e carico dei cluster

Sistemi Web geograficamente distribuiti



Web multi-cluster (2 livelli)

- Un unico hostname per il sito
- Un indirizzo IP per cluster



Dispatching di primo livello (mediante DNS)

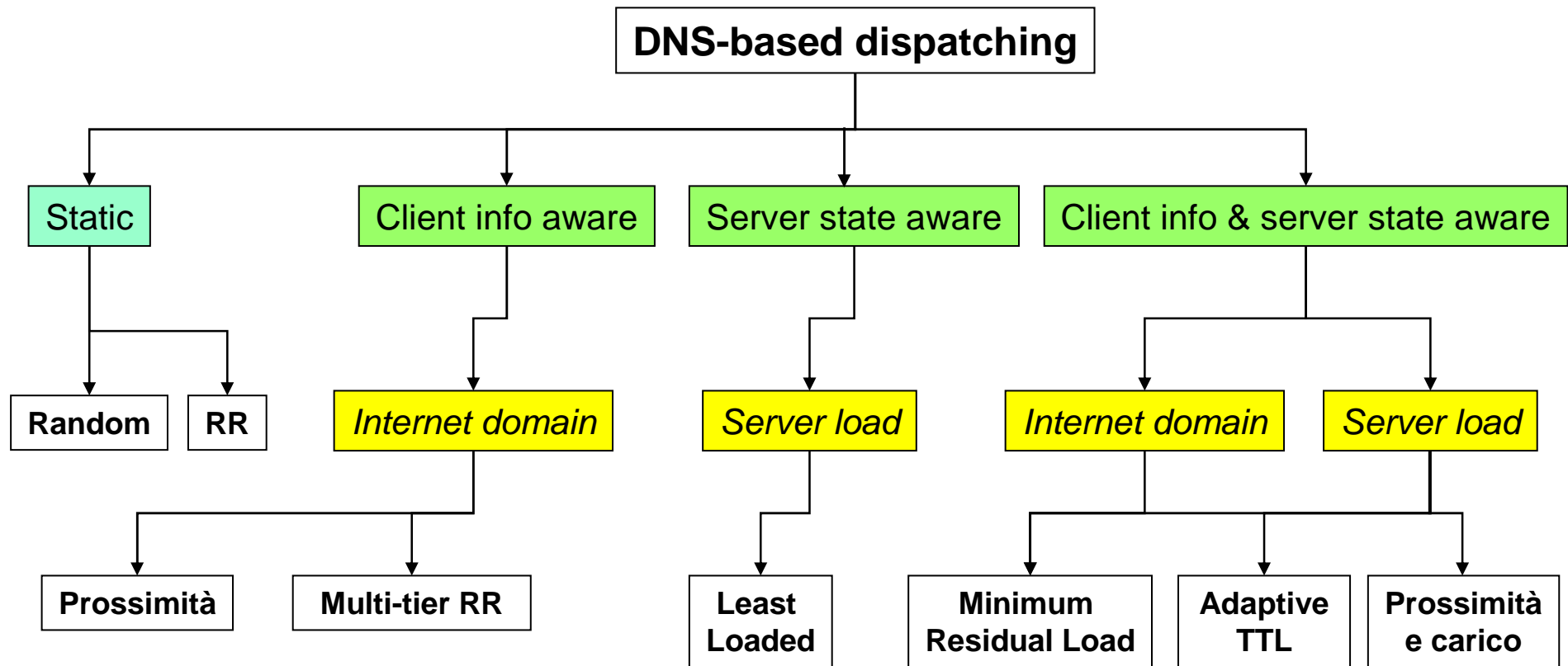
- Il primo livello di distribuzione geografico avviene nella fase di *risoluzione dell'indirizzo* (address lookup):
 - il client richiede l'**indirizzo IP** del cluster corrispondente all'hostname indicato nell'URL
 - se l'hostname è valido, il client riceve la coppia

< Indirizzo IP, Time-To-Live >

da:

- cache di qualche name server intermedio
- **DNS autoritativo del sito**, opportunamente modificato (integrato o meno da altro componente). Può applicare diverse politiche di dispatching per selezionare il Web cluster “migliore”

Algoritmi di dispatching per DNS



Prossimità in Internet

- La prossimità in Internet è un problema interessante: la prossimità geografica tra client e server **non implica** *prossimità Internet (round trip latency)*
- Valutazione *statica* della prossimità
 - indirizzo IP del client per determinare la zona Internet (simile a distanza geografica)
 - numero di hop (informazione “stabile” più che “statica”)
 - *network hops* (e.g., `traceroute`)
 - *Autonomous System hops* (query delle tabelle di routing)

Non garantisce la selezione del cluster migliore,
e.g., “*links are not created equal*”

Prossimità in Internet (2)

- Valutazione *dinamica* della prossimità
 - round trip time (es., `ping`, `tcpping`)
 - bandwidth disponibile (es., `cprobe`)
 - latenza delle richieste HTTP (es., *request emulation*)

Tempo aggiuntivo e costi di traffico per la valutazione

- Un problema ancora aperto: *correlazione tra il numero di hop ed il round trip time?*
 - “Vecchie” misure: prossima a zero
 - “Recenti” misure: elevata, mediamente elevata

Problemi del dispatching geografico

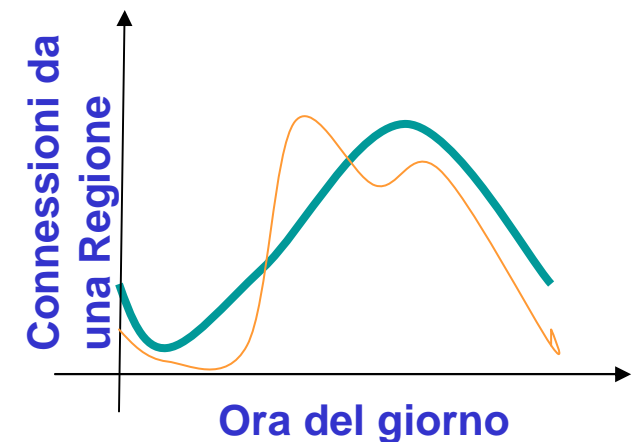
(di cui le politiche di dispatching devono tener conto)

Tipici problemi del dispatching Web

- Picchi di carico in alcune ore/giorni

Problemi aggiuntivi

- Traffico dipendente dai fusi orari
- Distribuzione non uniforme dei client tra le regioni Internet
- Prossimità Internet tra client e Web cluster
- (Per DNS) Caching di [hostname-indirizzo IP] in name server intermedi per l'intervallo del Time-To-Live



Problemi del dispatching DNS

- A causa del caching, nel caso di siti Web molto popolari, il DNS controlla solo il 5% del traffico in arrivo al sito
- A differenza del Web switch (che controlla il 100% del traffico in arrivo al sito), il DNS deve utilizzare algoritmi sofisticati (es., *TTL-adattativi*)
- Non sono stati trovati (*esistono?*) algoritmi di dispatching DNS in grado di evitare episodi di sovraccarico per tutte le classi di workload

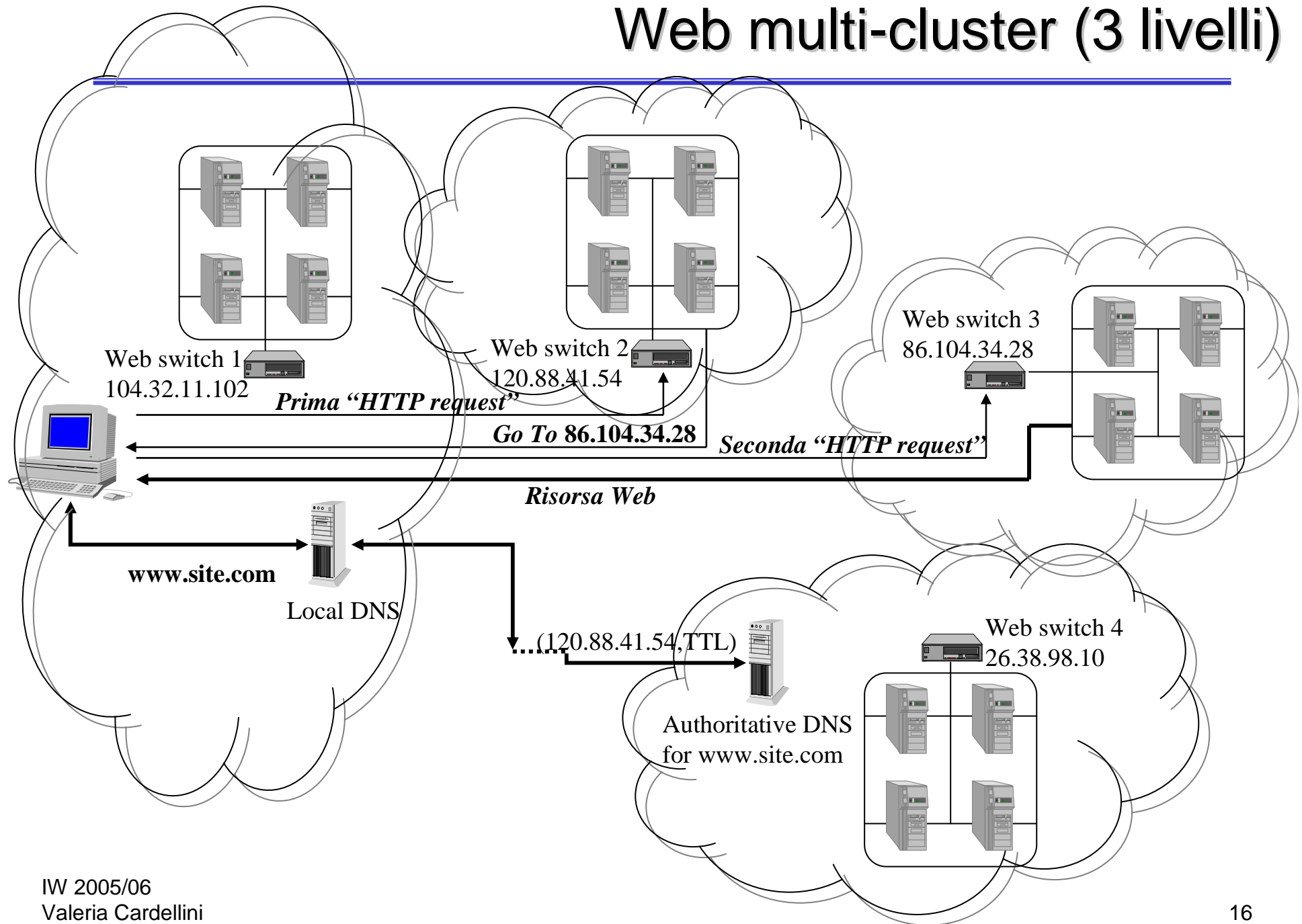
Come risolvere i problemi del dispatching DNS

- Fase di lookup: integrare il dispatching DNS con un'altra entità
- Fase di richiesta: integrare il dispatching DNS centralizzato con dispatching distribuito da parte dei server; alcuni meccanismi di re-routing delle richieste:
 - Ridirezione HTTP
 - IP tunneling
 - URL rewriting

Dispatching per Web multi-cluster

- Indirizzi del sito Web
 - Un unico hostname (e.g., “www.site.com”)
 - Un indirizzo IP per ogni Web cluster
- Tre livelli di routing e dispatching:
 1. **DNS** autoritativo seleziona il “Web cluster migliore”
 2. **Web switch** del cluster seleziona il “Web server migliore”
 3. Ciascun **Web server** (o Web switch di livello 7) può ridirigere le richieste verso un altro Web cluster (ad esempio, per risolvere situazioni temporanee di sovraccarico)

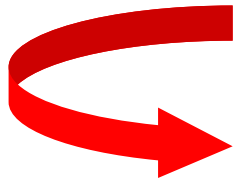
Web multi-cluster (3 livelli)



Motivazioni per terzo livello di dispatching

Web multi-cluster con **due** livelli di dispatching

- Controllo elevato sul carico che raggiunge il Web cluster (*buon bilanciamento intra-cluster*)
- Reazione lenta ad un cluster sovraccarico (*cattivo bilanciamento inter-cluster*)



Web multi-cluster con **tre** livelli di dispatching:
Reazione immediata per spostare il carico da un Web cluster sovraccarico
(meglio “HTTP redirection” di “IP tunneling”)

Ridirezione HTTP

- Il meccanismo di ridirezione è parte del protocollo HTTP ed è supportato dagli attuali browser
- **DNS e Web switch:** usano politiche di dispatching *centralizzate*
- **Ridirezione:** usa politiche di dispatching *distribuite*, in cui tutti i server Web possono partecipare al (ri-)assegnamento delle richieste
- La ridirezione è completamente trasparente per l'utente (non per il client!)

message header
HTTP status code
302 - "Moved temporarily" to a new location

- "New location"
 - ridirezione ad un indirizzo IP (**migliori prestazioni**)
 - ridirezione ad un hostname

Ridirezione HTTP: pro e contro

PRO

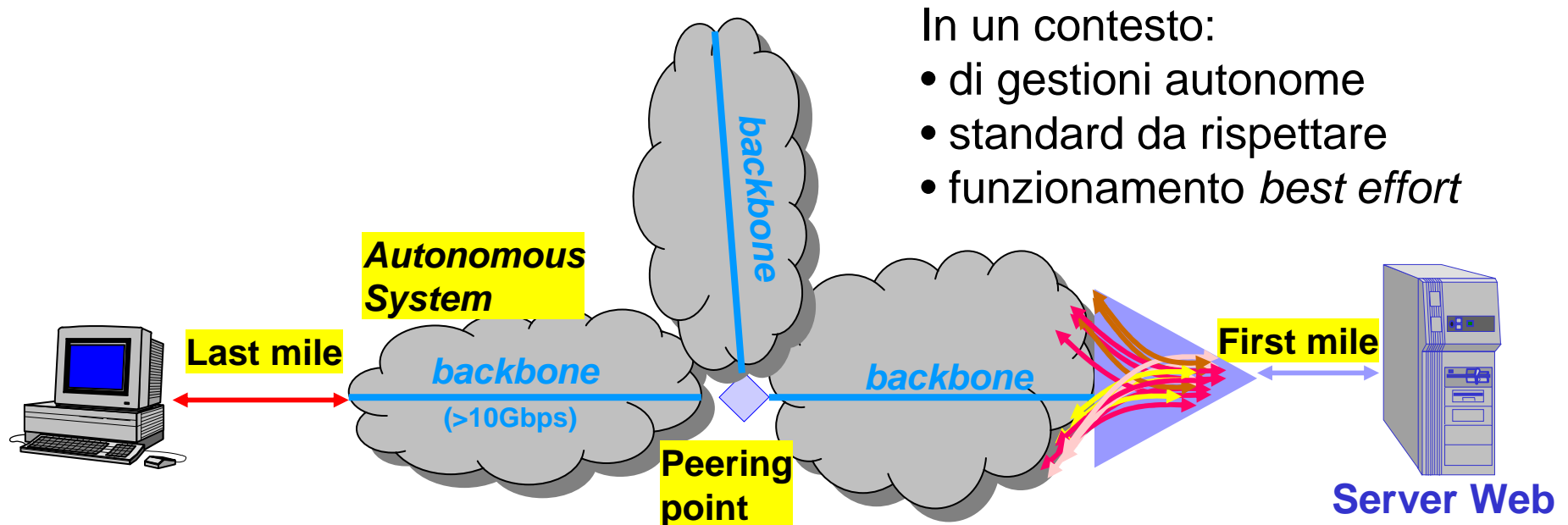
- Compatibile con tutti i client e server Web (implementata a livello applicativo)
- Meccanismo distribuito che soddisfa requisiti di affidabilità (non introduce “single point of failure”)
- Distribuzione delle richieste *content-aware*

CONTRO

- Limitata alla ridirezione di richieste HTTP (meccanismi di ridirezione più generali, ad es. IP tunneling)
- Aumenta il traffico in quanto ogni richiesta ridiretta richiede una nuova connessione TCP

Tuttavia, la ridirezione **riduce** il tempo di risposta quando
impatto del server > impatto della rete

Colli di bottiglia del delivery



- First mile:
 - linea di collegamento tra il server Web ed il resto di Internet
- Peering point
 - punto di connessione tra diversi sistemi autonomi
- Backbone
- Last mile
 - collegamento del client al proprio ISP

Content delivery

- **Content delivery**: strategia per distribuire il contenuto mediante rete IP su una base di pay-per-delivery
- Content Delivery Network o Content Distribution Network (CDN)
- Il content delivery ha molte probabilità di dar luogo a profitti

- Il problema del network delivery
 - Più di 9000 Autonomous System (AS)
 - Nessuna rete gestisce più del 5% del traffico
 - La grande maggioranza degli AS gestisce meno dell'1% del traffico
- Soluzione al problema del network delivery
 - Distribuire i server in molti (quasi tutti) gli AS

Content Delivery Network

- Content Delivery Network (o Content Distribution Network):
 - formata da un'infrastruttura di server (detti *content server*, *delivery server* o *edge server*) che lavorano insieme
 - i content server della CDN sono distribuiti su una vasta area geografica per permettere la fornitura del contenuto Web da locazioni più vicine all'utente
 - il content provider delega il servizio del contenuto del proprio sito Web (fornito dall'*origin server*) ad una compagnia che gestisce una CDN (*content outsourcing*)
 - i content server della CDN forniscono soltanto il contenuto dei siti Web gestiti dalla CDN

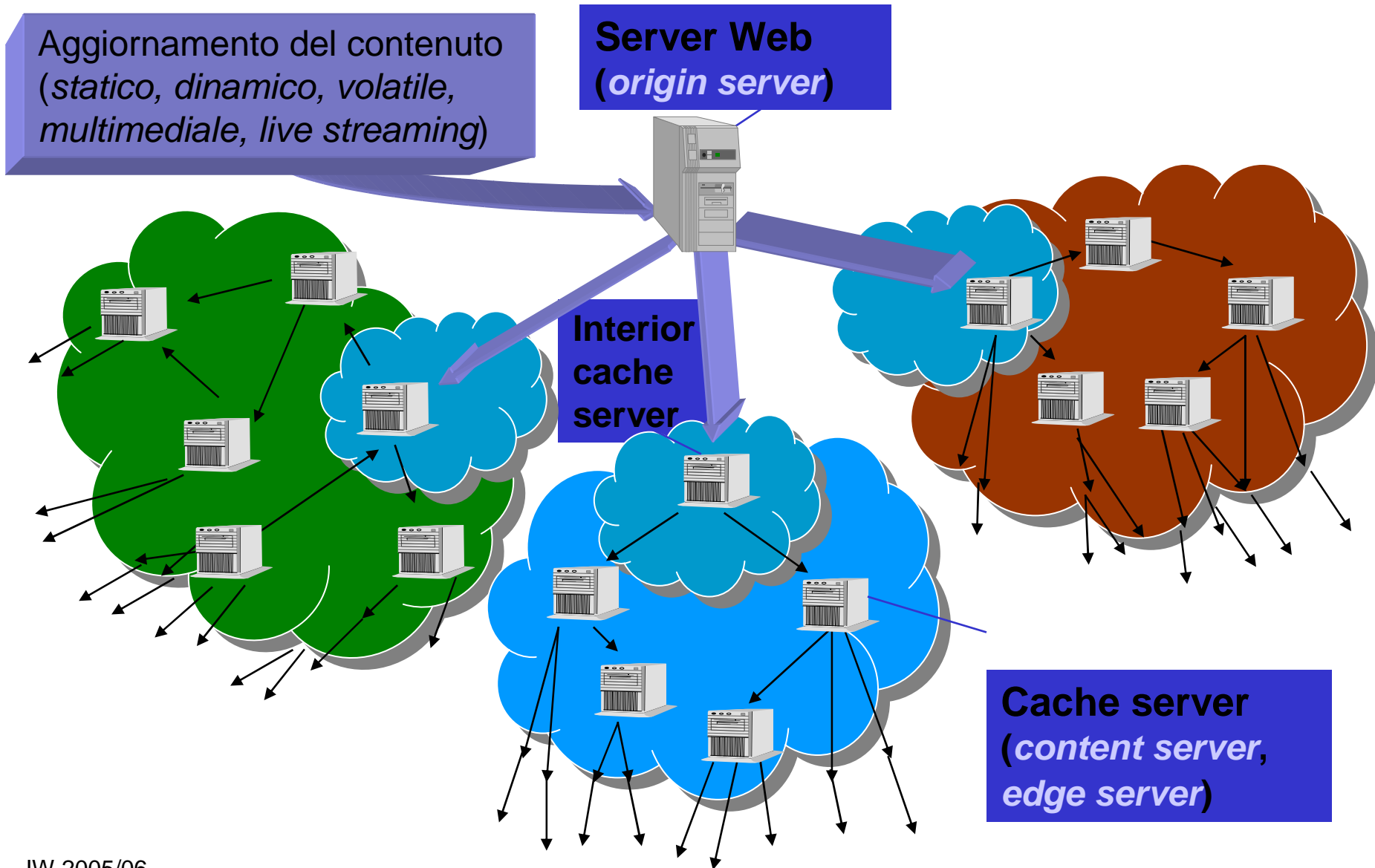
Web caching e CDN

- Web caching e CDN hanno in comune:
 - architettura costituita da un'infrastruttura di cache server che operano in modo cooperativo
 - i cache server sono distribuiti su una vasta area geografica per permettere la distribuzione del contenuto Web da locazioni più vicine all'utente

- Il Web caching opera in modo indipendente dai siti Web
- Costo per ISP (risparmio di banda), non per i content provider
- Caching generalizzato: viene memorizzato il contenuto di tutti i siti Web

- Il content provider delega la distribuzione del contenuto alla CDN
- Costo per i content provider, non per ISP
- Caching selettivo: viene memorizzato solo il contenuto dei siti Web gestiti dalla CDN

Architettura di una CDN



Perché usare una CDN?

- Lo scopo è analogo al Web caching:
 - avvicinare il contenuto del sito Web all'utente per ridurre il tempo di latenza ed il carico sul server Web originale
- I Web cluster possono essere lontani dagli utenti
- La congestione di Internet può far fallire le migliori architetture di server Web
- I picchi di traffico (*flash crowds*) possono provocare un crash dei server Web
- I sistemi Web distribuiti geograficamente richiedono che il content provider posseda l'infrastruttura di servizio

Vantaggi di CDN rispetto a Web caching

Web caching

- cache dei server proxy riempite in seguito alle richieste degli utenti
- overhead sul sito Web derivante dall'offloading ripetuto per ogni proxy server che accede al sito
- il content provider non ha il controllo del contenuto nella cache dei proxy server (*elevati problemi di consistenza*)
- il content provider perde molte informazioni sulla distribuzione del proprio contenuto

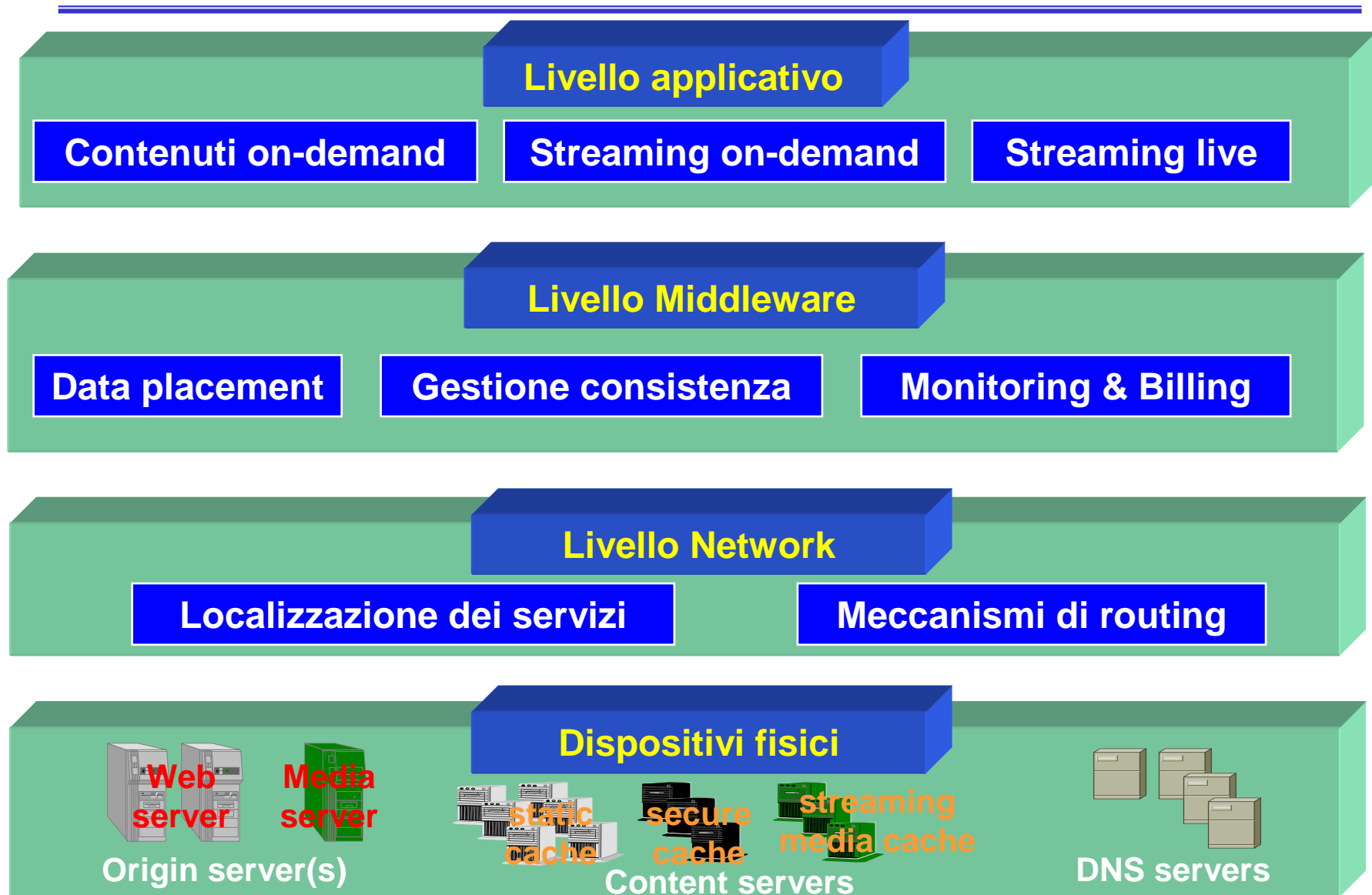
CDN

- contenuto memorizzato in anticipo (*prefetching*)
- overhead sul sito Web derivante dall'unico offloading
- il content provider ha il controllo diretto del contenuto dei content server
- una CDN può fornire al content provider statistiche accurate sui pattern di accesso al sito Web

Contenuti serviti da CDN

- Una CDN è generalmente utilizzata per fornire i seguenti servizi di delivery:
 - Distribuzione **contenuti non streaming**
 - Contenuto statico delle pagine Web, in particolare immagini
 - Contenuti dinamico (in parte)
 - Ambiguità nel termine “dinamico”: in realtà, si intende “volatile” (contenuto statico che cambia frequentemente)
 - Contenuto con autenticazione
 - Solo autenticazione prevista dal protocollo HTTP
 - Distribuzione **contenuti streaming on-demand**
 - Il contenuto è digitalizzato e memorizzato come media file su media server. Esempio: video-on-demand, clip musicali
 - Distribuzione **contenuti streaming live**
 - Il contenuto è distribuito quasi istantaneamente come media file. Esempio: eventi sportivi e musicali

Componenti dell'architettura di una CDN



Meccanismi di routing per CDN

come indirizzare la richiesta al server “migliore”

- Una questione fondamentale riguarda il meccanismo di routing con cui ridirigere una richiesta di un client per un oggetto servito dalla CDN verso un content server
- Principali tecniche adottate:
 - uso del Domain Name System: **DNS redirection**
 - riscrittura URL con hostname simbolici: **URL rewriting**
 - riscrittura URL con indirizzo IP: **URL-IP rewriting**
- Altre possibilità
 - Redirezione mediante protocollo HTTP/RTSP
 - Anycast, IP tunneling
- Le società CDN usano nomi altisonanti (es. *Global Traffic Management system*), ma in realtà i principali dispositivi per il routing sono costituiti da **DNS autoritativi** arricchiti con funzionalità di controllo sullo stato della rete e dei content server

DNS redirection

- Il **DNS autoritativo** del sito Web delega la risoluzione dell'hostname della URL in un indirizzo IP a un **DNS controllato dalla CDN**. Nella scelta da hostname a indirizzo IP si può effettuare la scelta di un determinato content server
- Il valore del Time-To-Live (TTL) assegnato dal DNS autoritativo è prossimo a 0, in modo tale che la CDN possa modificare spesso il mapping tra hostname e indirizzo IP per:
 - facilitare il bilanciamento del carico tra i content server
 - limitare il caching delle risoluzioni indirizzi

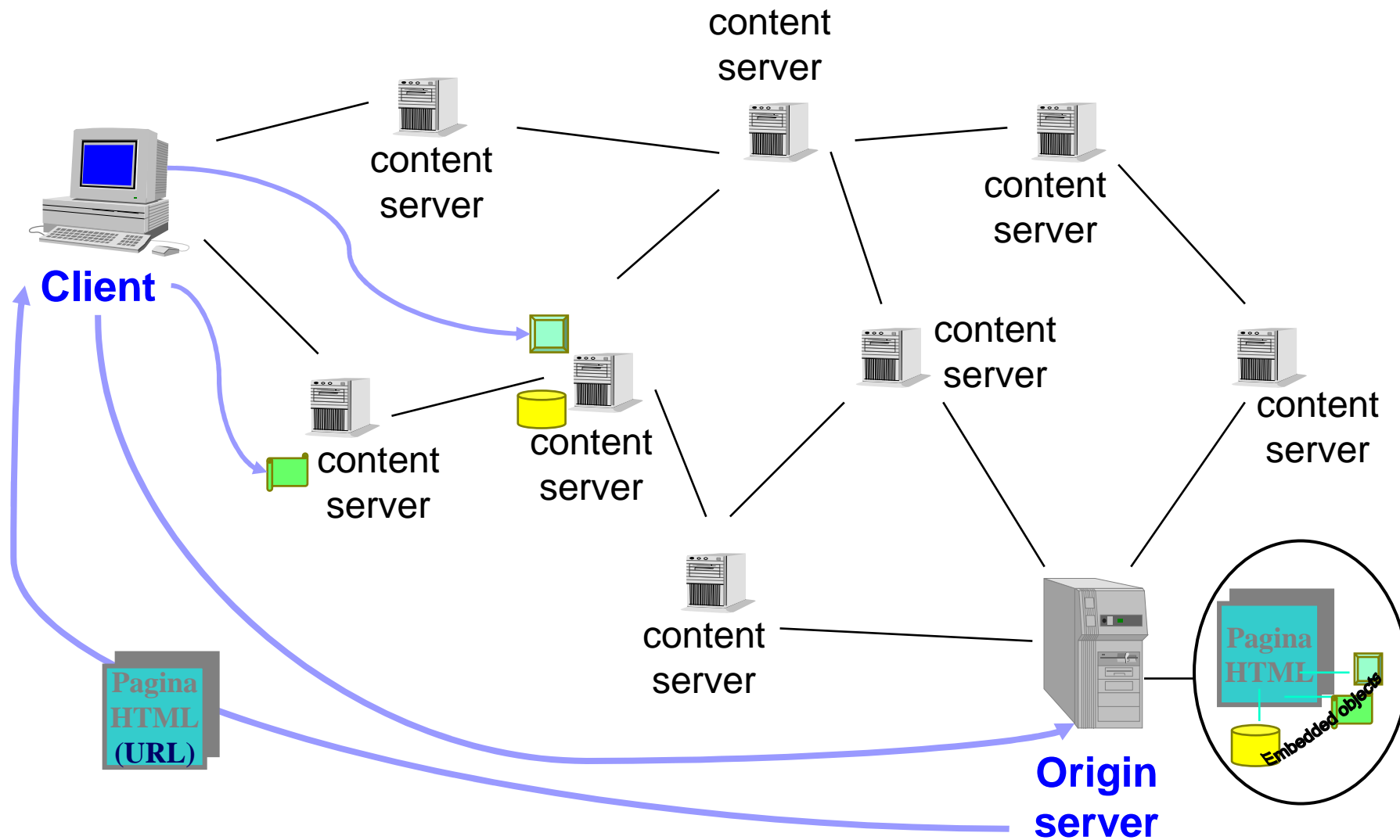
DNS redirection (2)

- Esistono due tipi di CDN che adottano la tecnica del routing mediante DNS:
 - *full-content delivery* (o *first hit at CDN*)
 - L'origin server è nascosto a tutti, eccetto che alla CDN
 - L'origin server modifica il proprio server DNS autoritativo in modo da inviare le richieste di mapping tra il nome e l'indirizzo IP al server autoritativo della CDN (*DNS outsourcing*)
 - tutte le richieste indirizzate all'origin server arrivano ai content server
 - *partial-content delivery* (o *first hit at origin*)
 - L'origin server modifica nella pagina HTML le URL degli oggetti incorporati nella pagina (soprattutto immagini), in modo tale che siano risolti dal name server autoritativo della CDN (ad esempio, www.foo.com/bar.gif viene modificato in cdn-foo.net/www.foo.com/bar.gif)

URL rewriting

- L'origin server riscrive dinamicamente le URL presenti nella pagina HTML base in modo da ridirigere le richieste dei client per gli *embedded object* verso un determinato content server della CDN
- Due alternative di **rewriting**
 - URL riscritta **con indirizzo IP**: indirizzamento diretto del content server
 - URL riscritta **con hostname**: indirizzamento indiretto del content server
 - Nella fase di risoluzione del nuovo hostname, i DNS della CDN selezioneranno il content server opportuno

Routing mediante URL rewriting



Diffusione delle CDN

- Alcune società di CDN:
 - Akamai <http://www.akamai.com>
 - Mirror Image <http://www.mirror-image.com>
 - Speedera <http://www.speedera.com>
- Diffusione molto rapida dell'uso di CDN:
 - Nel Novembre 1999, soltanto 1-2% dei 700 siti Web più popolari utilizzavano CDN per servire il proprio contenuto
 - Nel Dicembre 2000, il 17% dei 1030 siti Web più popolari utilizzavano CDN per servire il proprio contenuto
 - Nel Dicembre 2000 tra i 39 siti Web più popolari che usavano CDN, 37 utilizzavano Akamai
 - Nel Dicembre 2000, gli oggetti serviti dalle CDN erano immagini nel 96-98% dei casi

Soluzione di Akamai

- Un sito Web che vuole avere parte delle proprie risorse distribuita da Akamai deve rinominare le URL ad esse relative con un prefisso specifico
 - il prefisso include il nome di un host, ad esempio
a799.g.akamai.net
- Risoluzione del nome in un indirizzo IP di un content server di Akamai eseguita dal server DNS di Akamai
- Il content server prescelto è “vicino” al name server locale del client

Soluzione di Akamai (2)

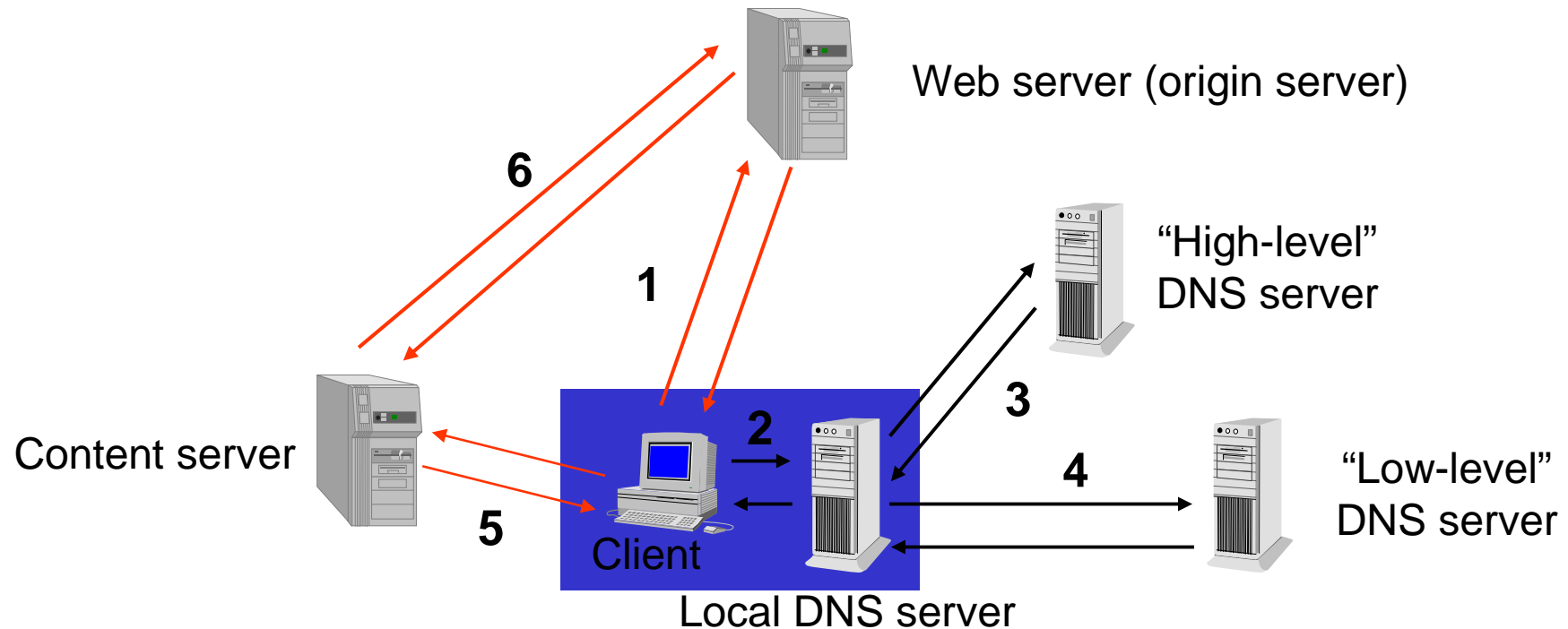
- Esempio:
 - richiesta per la pagina Web <http://www.xyz.com/> che contiene la risorsa [image.gif](#)
 - l'URL della risorsa <http://www.xyz.com/image.gif> è rinominata come <http://a799.g.akamai.net/3/799/388/9fd0a26b9d5686/www.xyz.com/image.gif>
 - il prefisso [a799.g.akamai.net](#) individua un content server di Akamai che possiede la risorsa
 - tramite il DNS, il client risolve [a799.g.akamai.net](#) in un indirizzo IP, ad esempio [193.206.138.7](#) (distante solo 5 hop dal client, anziché più di 17 come gli origin server di [www.msnbc.com](#)!)

Soluzione di Akamai (3)

- Esempio (*continua*):
 - Se il content server 193.206.138.7 possiede la risorsa image.gif, la invia al client
 - Altrimenti, la richiede tramite un protocollo interno ad un altro content server di Akamai oppure al server originale www.xyz.com e la memorizza nella propria cache
 - Un altro client appartenente ad un'altra rete che richiede la stessa pagina avrà un link uguale per l'immagine
[http://a799.g.akamai.net/3/799/388/9fd0a26b9d5686/
www.xyz.com/image.gif](http://a799.g.akamai.net/3/799/388/9fd0a26b9d5686/www.xyz.com/image.gif) ma l'indirizzo IP corrispondente potrebbe essere diverso

DNS redirection in Akamai

- Akamai utilizza un doppio livello di risoluzione dei nomi attuato dal DNS
 - Esempio (dopo la prima risoluzione del DNS)



Un esempio: il sito di Torino 2006

- Il sito di Torino 2006 (<http://www.torino2006.org>) è gestito da Akamai con un approccio di tipo **full-content delivery**
- Come scoprirlo? Usiamo il comando `dig`

```
claudius:~$ dig www.torino2006.org
; <<>> DiG 9.2.5 <<>> www.torino2006.org
...
;; QUESTION SECTION:
;www.torino2006.org.      IN      A

;; ANSWER SECTION:
www.torino2006.org.     77677  IN      CNAME  torino2006.org.edgesuite.net.
torino2006.org.edgesuite.net. 12877  IN      CNAME  a835.g.akamai.net.
a835.g.akamai.net.     20     IN      A      80.239.170.222
a835.g.akamai.net.     20     IN      A      80.239.170.217
```

TTL pari a 20 sec.

Web cluster

Un esempio: il sito di Torino 2006 (2)

;; AUTHORITY SECTION:

g.akamai.net.	1069	IN	NS	n6g.akamai.net.
g.akamai.net.	1069	IN	NS	n7g.akamai.net.
g.akamai.net.	1069	IN	NS	n8g.akamai.net.
g.akamai.net.	1069	IN	NS	n0g.akamai.net.
g.akamai.net.	1069	IN	NS	n1g.akamai.net.
g.akamai.net.	1069	IN	NS	n2g.akamai.net.
g.akamai.net.	1069	IN	NS	n3g.akamai.net.
g.akamai.net.	1069	IN	NS	n4g.akamai.net.
g.akamai.net.	1069	IN	NS	n5g.akamai.net.

I name server
di Akamai

;; ADDITIONAL SECTION:

n0g.akamai.net.	795	IN	A	195.43.191.23
n1g.akamai.net.	2594	IN	A	195.43.191.25
n2g.akamai.net.	2581	IN	A	195.43.191.30
n3g.akamai.net.	795	IN	A	195.43.191.31
n4g.akamai.net.	795	IN	A	195.43.191.23
n5g.akamai.net.	2594	IN	A	195.43.191.25
n6g.akamai.net.	795	IN	A	213.156.44.148
n7g.akamai.net.	2594	IN	A	63.210.47.50
n8g.akamai.net.	795	IN	A	213.156.44.148

Un esempio: il sito di Torino 2006 (3)

- Dove sono localizzati i cluster di indirizzo 80.239.170.222 e 80.239.170.217?
- Usiamo ad un servizio di traceroute grafico (ad es. VisualRoute, <http://www.visualroute.it/vr.asp>)
- Tutti e due i cluster sono localizzati in Francia
- Usando un altro Internet Service Provider (oppure risolvendo il nome del sito in un diverso momento) è possibile ottenere un risultato diverso per la risoluzione del nome `www.torino2006.org` in un indirizzo IP

Caching di contenuti dinamici

- Problema di ricerca aperto
 - *Active Cache*: applet Java eseguiti sui cache server
 - Soluzioni integrate con tecnologie di database, es. Oracle 10g
- Tecnologia **Edge Side Includes (ESI)**
 - <http://www.esi.org/>
 - E' un consorzio costituito da Akamai, Oracle, IBM, Mirror Image, BEA Systems, ATG, Vignette, ...
 - Linguaggio di markup per descrivere componenti di una pagina Web che sono cacheable e non-cacheable
 - Tag XML per specificare frammenti ESI in una pagina Web
 - Per ciascun frammento ESI si può specificare un diverso requisito per il caching

Tecnologia Edge Side Includes



Il campo TTL può essere usato per specificare quando la copia del frammento non è più valida

Content peering

- Cooperazione ed interoperabilità tra più CDN allo scopo di migliorare la scalabilità, la tolleranza ai guasti e le prestazioni
- Le soluzioni proposte per il content peering adottano il routing tra CDN effettuato a livello del Domain Name System
 - Il server DNS (*potenziato*) deve essere in grado di fornire l'indirizzo di una CDN che serve una regione geografica
 - L'intera CDN agisce come un content server, recuperando il contenuto e tariffando il costo dell'operazione alla CDN richiedente