

Reti di Calcolatori

Prof. Roberto Canonico

Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione

Corso di Laurea in Ingegneria Informatica

Content Delivery Networks (CDN)

**I lucidi presentati al corso sono uno strumento didattico
che NON sostituisce i testi indicati nel programma del corso**

Nota di copyright per le slide COMICS

Nota di Copyright

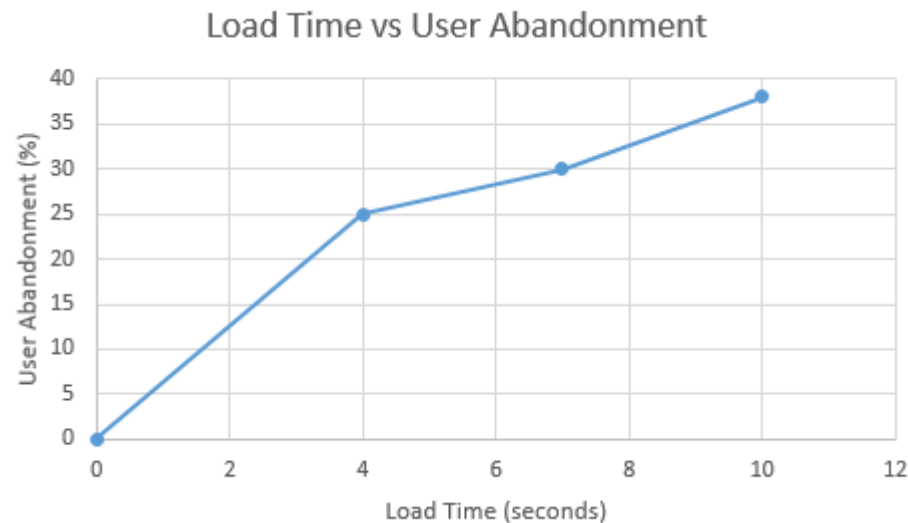
Questo insieme di trasparenze è stato ideato e realizzato dai ricercatori del Gruppo di Ricerca COMICS del Dipartimento di Informatica e Sistemistica dell'Università di Napoli Federico II. Esse possono essere impiegate liberamente per fini didattici esclusivamente senza fini di lucro, a meno di un esplicito consenso scritto degli Autori. Nell'uso dovranno essere esplicitamente riportati la fonte e gli Autori. Gli Autori non sono responsabili per eventuali imprecisioni contenute in tali trasparenze né per eventuali problemi, danni o malfunzionamenti derivanti dal loro uso o applicazione.

Autori:

Simon Pietro Romano, Antonio Pescapè, Stefano Avallone,
Marcello Esposito, Roberto Canonico, Giorgio Ventre

WWW: Importanza della consegna “timely”

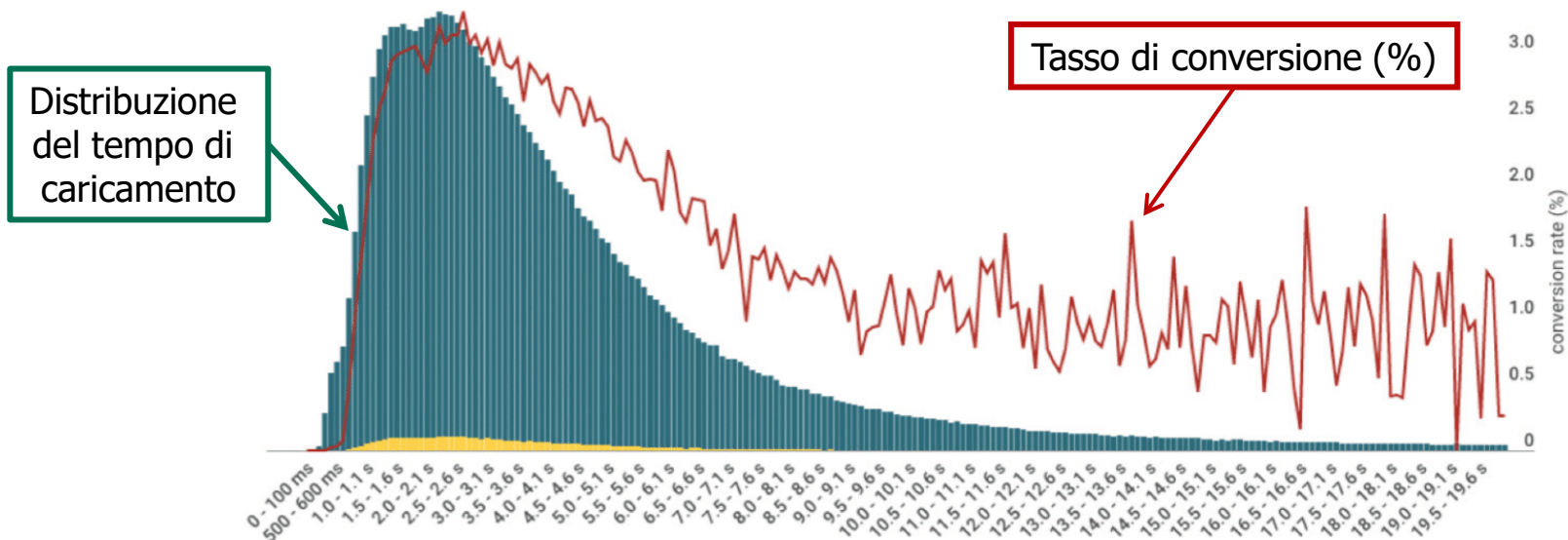
- Per un sito web una metrica importantissima è il TTI (*time to interact*)
 - TTI è il tempo che la pagina web impiega per rendere i suoi contenuti fruibili all'utente
- Una ricerca di Radware mostra che il 57% dei consumatori online abbandona una pagina web di un sito di e-commerce che richieda più di 3 secondi per caricarsi
- Se il tempo di caricamento di una pagina web supera i 10s, circa il 40% degli utenti rinuncia a proseguire la navigazione



Tasso di conversione (e-commerce)

- Il tasso di conversione (*conversion rate*) di un sito di e-commerce è la frazione percentuale di visite al sito che si traduce in una transazione commerciale (tipicamente, una vendita)
$$\text{Conversion Rate} = \text{Number of Sales} / \text{Number of Visits}$$
- Il tasso di conversione è fortemente influenzato dal tempo di caricamento delle pagine web del sito (particolarmente per l'accesso da terminali mobili)
 - Amazon stima che un aumento della latenza media di 100 ms comporti una perdita di ricavi dell'1%

How does load time correlate to conversion rate? (mobile)



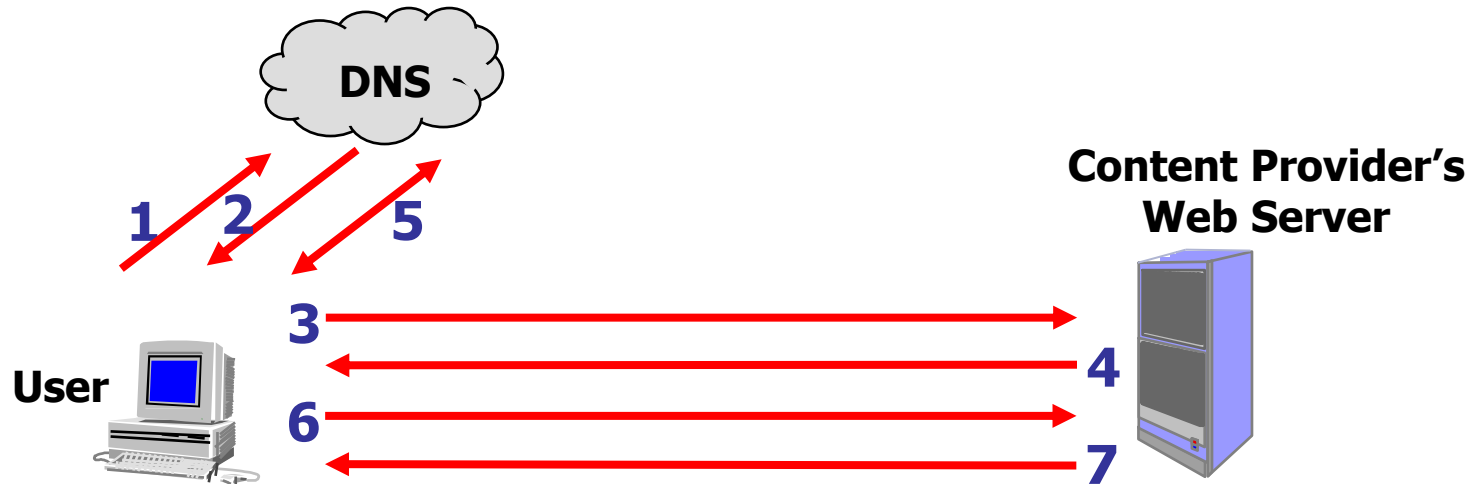
Cosa è una CDN

- Una Content Delivery Network è un'infrastruttura creata per distribuire efficacemente agli utenti di Internet i contenuti dei siti web più popolari
- Una CDN si basa sulla distribuzione di repliche dei contenuti dal server principale del “Content Provider” ad una molteplicità di server disposti sulla rete da un “Content Delivery Operator”
- Si presenta come un servizio a pagamento del quale usufruiscono i gestori dei siti web commerciali più popolari

Obiettivi di una CDN

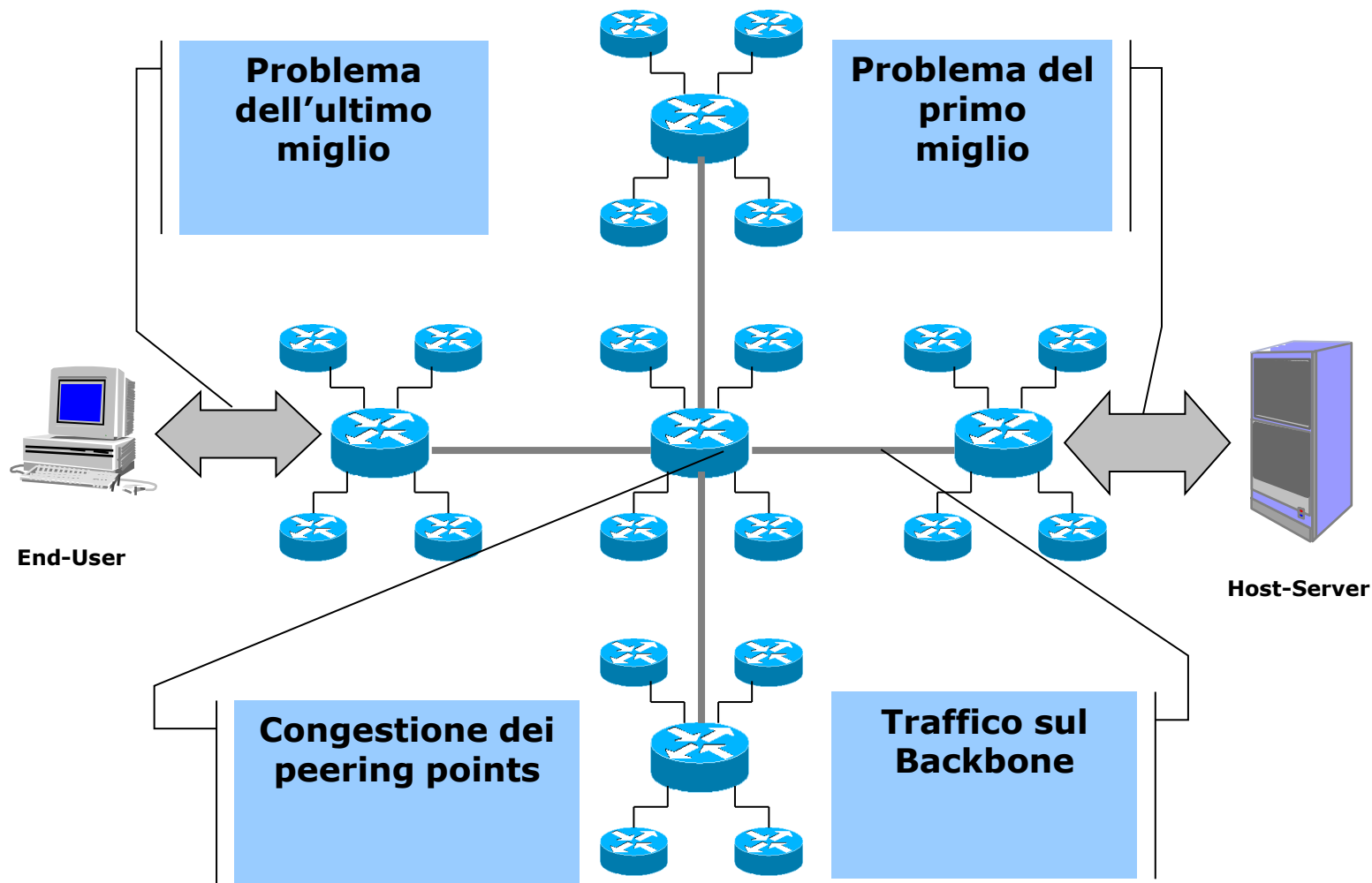
1. Alleviare il server web “master” dal carico degli utenti, in particolare proteggerlo da picchi di traffico improvvisi (flash crowds)
2. Offrire i contenuti ai singoli utenti tramite server collocati in prossimità degli utenti (alla periferia della rete)
3. Rendere il sistema di distribuzione dei contenuti più affidabile e robusto ai guasti

Interazioni che determinano il tempo di accesso ad una pagina web



- **User enters `www.xyz.com`**
 - **Browser requests IP address for `www.xyz.com`**
 - **DNS returns IP address**
- **Browser requests HTML**
- **Content provider's web server returns HTML**
- **Browser obtains IP addresses for hostnames listed in URLs of objects embedded on page**
 - **Browser requests embedded objects**
 - **Content provider's web server returns embedded objects**

Limiti dell'approccio centralizzato

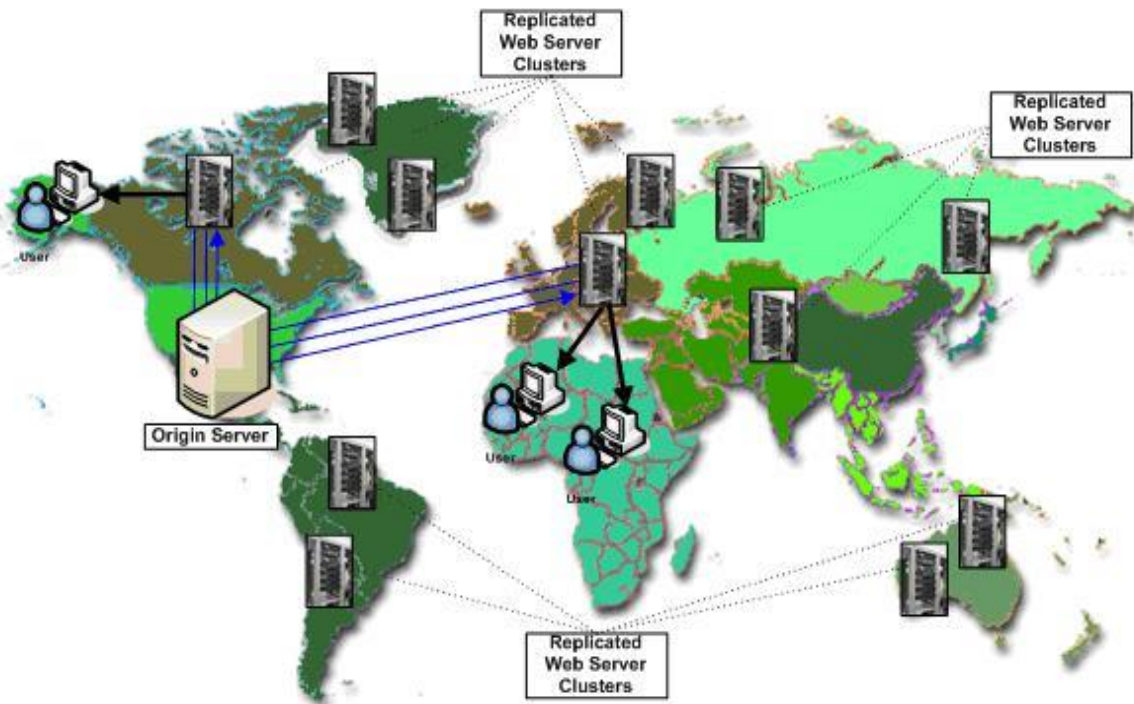


- Tramite una infrastruttura, spesso privata, distribuiscono, in maniera capillare i contenuti di uno specifico Content Provider
- Utilizzano forme proprietarie di caching basate su una complessa gestione del DNS, caratterizzata, tra l'altro, dalla conoscenza dell'indirizzo IP del Client
- Gestione centralizzata dei contenuti

Servizi offerti

- Le Content Delivery Network offrono ai Content Provider la possibilità di raggiungere, con una certa QoS, una vasta utenza
- Le CDN, d'altra parte, propongono a ISP di medie e grandi dimensioni, di collaborare, spesso gratuitamente, alla loro struttura

Content Delivery Networks



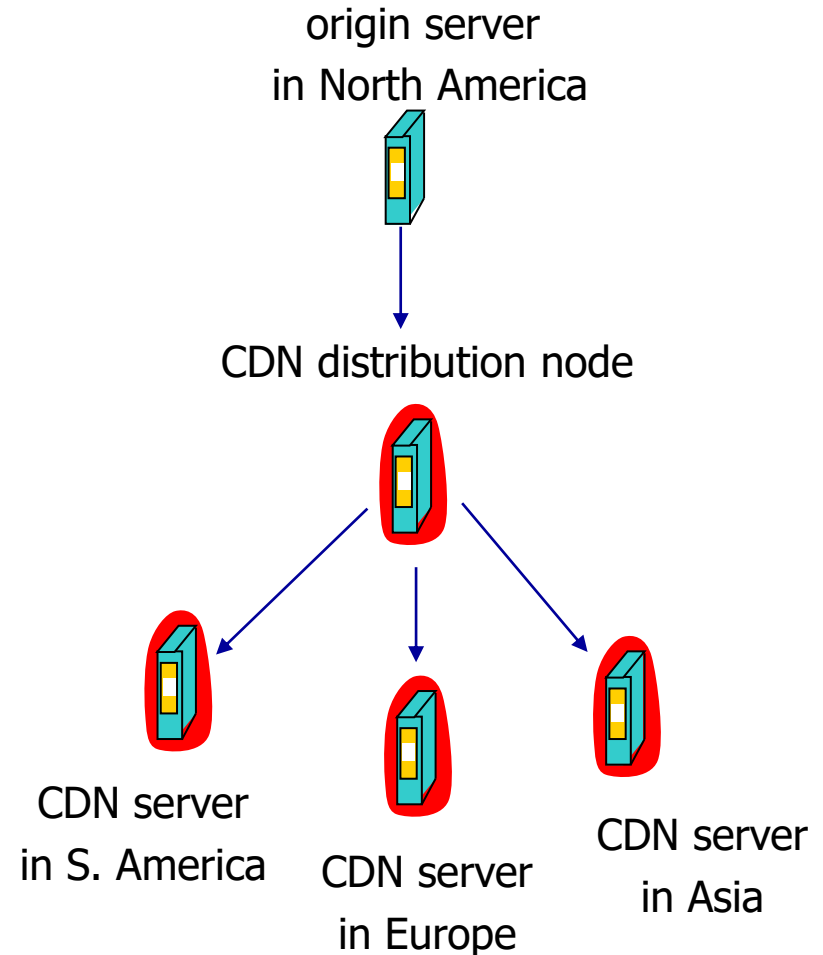
- **ottimizzano l'uso delle risorse di Internet avvicinando il contenuto all'utente**
- **benefici**
 - per gli utenti
 - per gli ISP
 - per i Content Provider
- **obiettivi**
 - **bassa latenza**
 - **bassi costi**
 - **raggiungibilità**
 - **protezione da flash events**

CDN: rilevanza del mercato ed attori

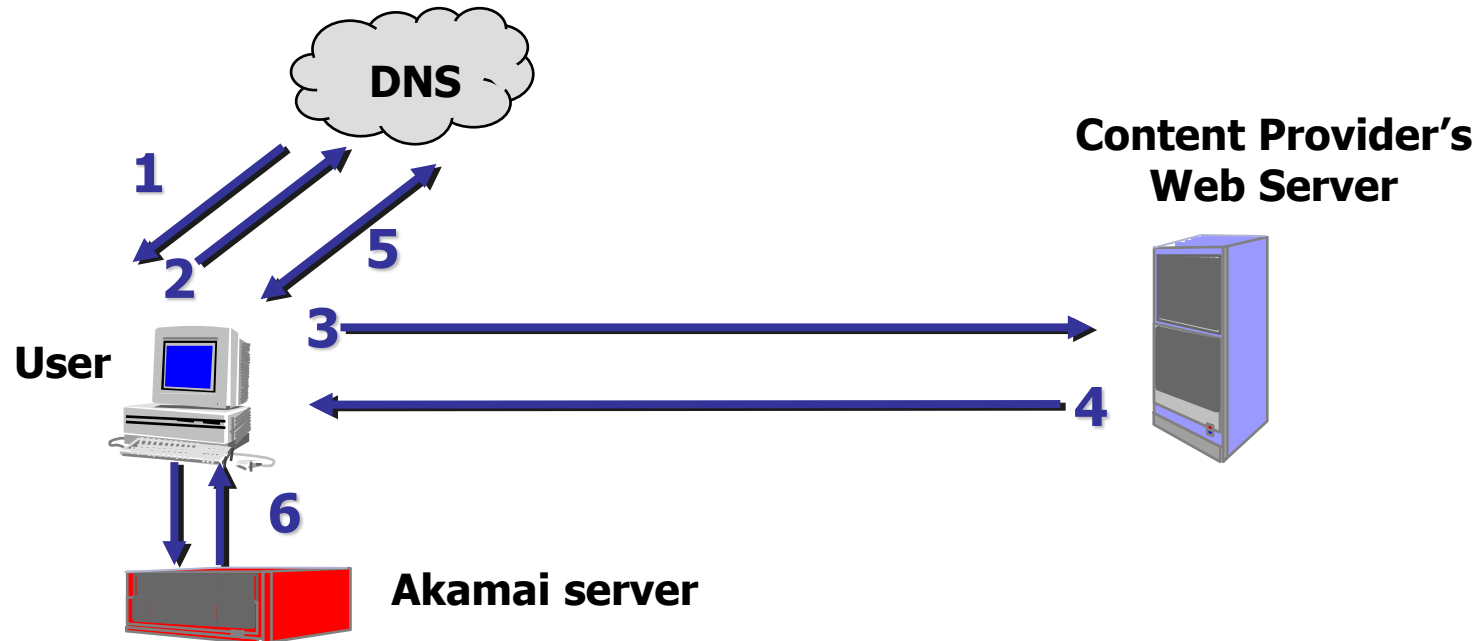
- In 2014, CDN market estimated approximately \$3.71 billion
- Market expected to grow to about \$12.16 billion by 2019
- Main players
 - **Akamai**
 - Founded in 1998 by MIT's researchers Tom Leighton e Danny Lewin
 - more than 240,000 servers in over 130 countries within 1,700 networks (as of 2018)
 - the most pervasive content delivery network
 - **Amazon Cloudfront**
 - delivers static and streaming contents
 - works seamlessly with other Amazon Cloud solutions
 - **CDNetworks** (world's third largest, and Asia's #1 full-service provider)
 - POPs in six continents, including 20 POPs in China
 - **Level3**
 - supports a comprehensive encoding suite for video, and intelligent traffic manager services
 - **Limelight**
 - 6,000 server at 75 POPs, and more than 30 regional content delivery centers in the United States, Europe, and Asia
 - **ChinaCache**
 - CDN market leader in China, with 127 POPs and 11,000 servers

Replica dei Contenuti

- Il customer della CDN (e.g., Akamai) è il content provider (e.g., CNN)
- La CDN replica il contenuto del customer sui propri CDN server
- Quando il provider (CNN) aggiorna il contenuto, la CDN aggiorna i propri server



Come funziona Akamai

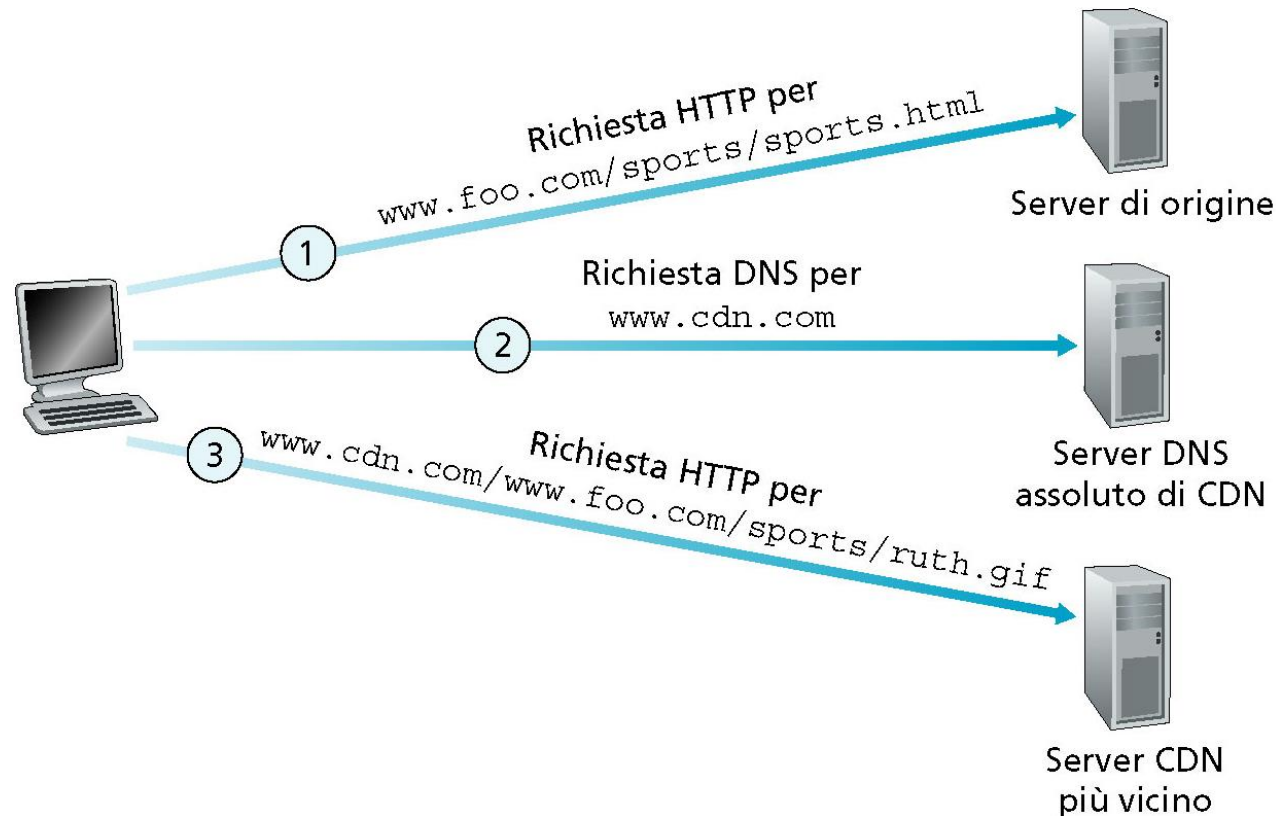


- User enters `www.xyz.com` and browser requests IP address for `www.xyz.com`
- DNS returns IP address
- Browser requests HTML
- Content provider's web server returns page with *Akamaized* URLs
- Browser obtains IP address of optimal Akamai server for embedded objects
- Browser obtains objects from optimal Akamai server

Prelievo da parte del client

Dal server origine al nodo CDN: routing delle richieste

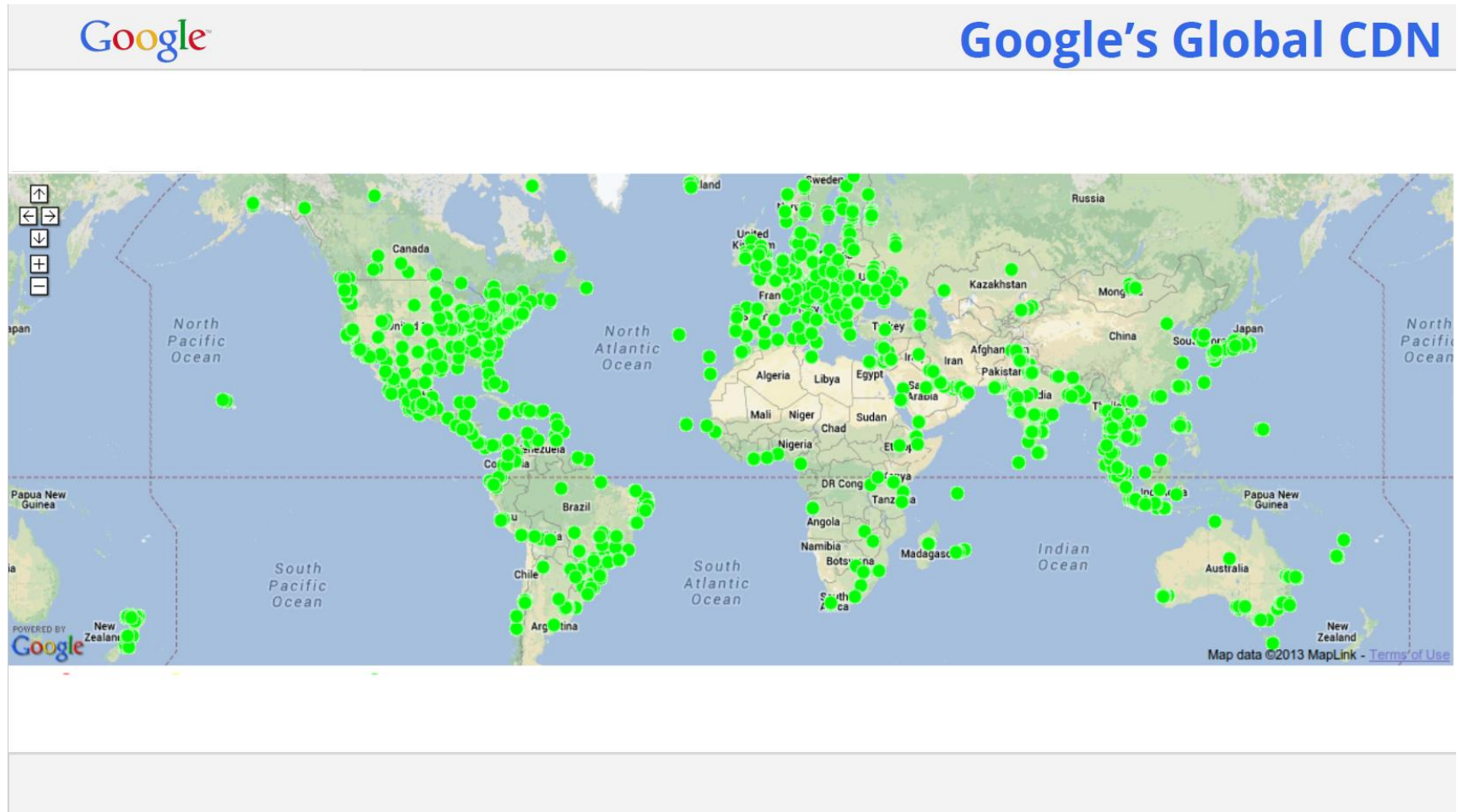
- La CDN crea una mappa che indica le distanze tra i vari ISP e i nodi CDN
- Quando arriva una query al DNS aut.
 - Si determina l'ISP che ha originato la query
 - Si usa la mappa per la scelta del server CDN più vicino



Google AMP

- Google AMP (Accelerated Mobile Pages Project) è un progetto reso pubblico ad ottobre 2015
- Si basa su un framework basato su tecnologie web esistenti che consente di creare pagine web leggere visualizzabili rapidamente da dispositivi mobili
- Il progetto AMP si basa su tre componenti principali:
 - AMP HTML: un HTML con alcune restrizioni
 - AMP JS: una libreria che consente l'esecuzione di codice JavaScript con delle limitazioni al fine di ottenere un rendering veloce delle pagine
 - AMP Cache: un sistema di web cache che può essere usato per servire pagine AMP HTML precedentemente memorizzate
- Attualmente, ci sono due provider AMP Cache:
 - Google AMP Cache
 - Cache AMP Cloudflare

La CDN di Google



E' l'infrastruttura su cui Google ha costruito il sistema AMP Cache