Multi-layer switch
commutazione hardware a vari livelli

Mario Baldi

Politecnico di Torino

http://staff.polito.it/mario.baldi

Basato sul capitolo 10 di:
M. Baldi, P. Nicoletti, "Switched LAN", McGraw-Hill, 2002, ISBN 88-386-3426-2

Nota di Copyright

Questo insieme di trasparenze (detto nel seguito slide) è protetto dalle leggi sul copyright e dalle disposizioni dei trattati internazionali. Il titolo ed i copyright relativi alle slide (ivi inclusi, ma non limitatamente, ogni immagine, fotografia, animazione, video, audio, musica e testo) sono di proprietà degli autori indicati a pag. 1.

Le slide possono essere riprodotte ed utilizzate liberamente dagli istituti di ricerca, scolastici ed universitari afferenti al Ministero dell'Istruzione, dell'Università e della Ricerca, per scopi istituzionali, non a fine di lucro. In tal caso non è richiesta alcuna autorizzazione.

Ogni altra utilizzazione o riproduzione (ivi incluse, ma non limitatamente, le riproduzioni su supporti magnetici, su reti di calcolatori e stampate) in toto o in parte è vietata, se non esplicitamente autorizzata per iscritto, a priori, da parte degli autori.

L'informazione contenuta in queste slide è ritenuta essere accurata alla data della pubblicazione. Essa è fornita per scopi meramente didattici e non per essere utilizzata in progetti di impianti, prodotti, reti, ecc. In ogni caso essa è soggetta a cambiamenti senza preavviso. Gli autori non assumono alcuna responsabilità per il contenuto di queste slide (ivi incluse, ma non limitatamente, la correttezza, completezza, applicabilità, aggiornamento dell'informazione).

In ogni caso non può essere dichiarata conformità all'informazione contenuta in queste slide.

In ogni caso questa nota di copyright non deve mai essere rimossa e deve essere riportata anche in utilizzi parziali.

Layer 4 Switch

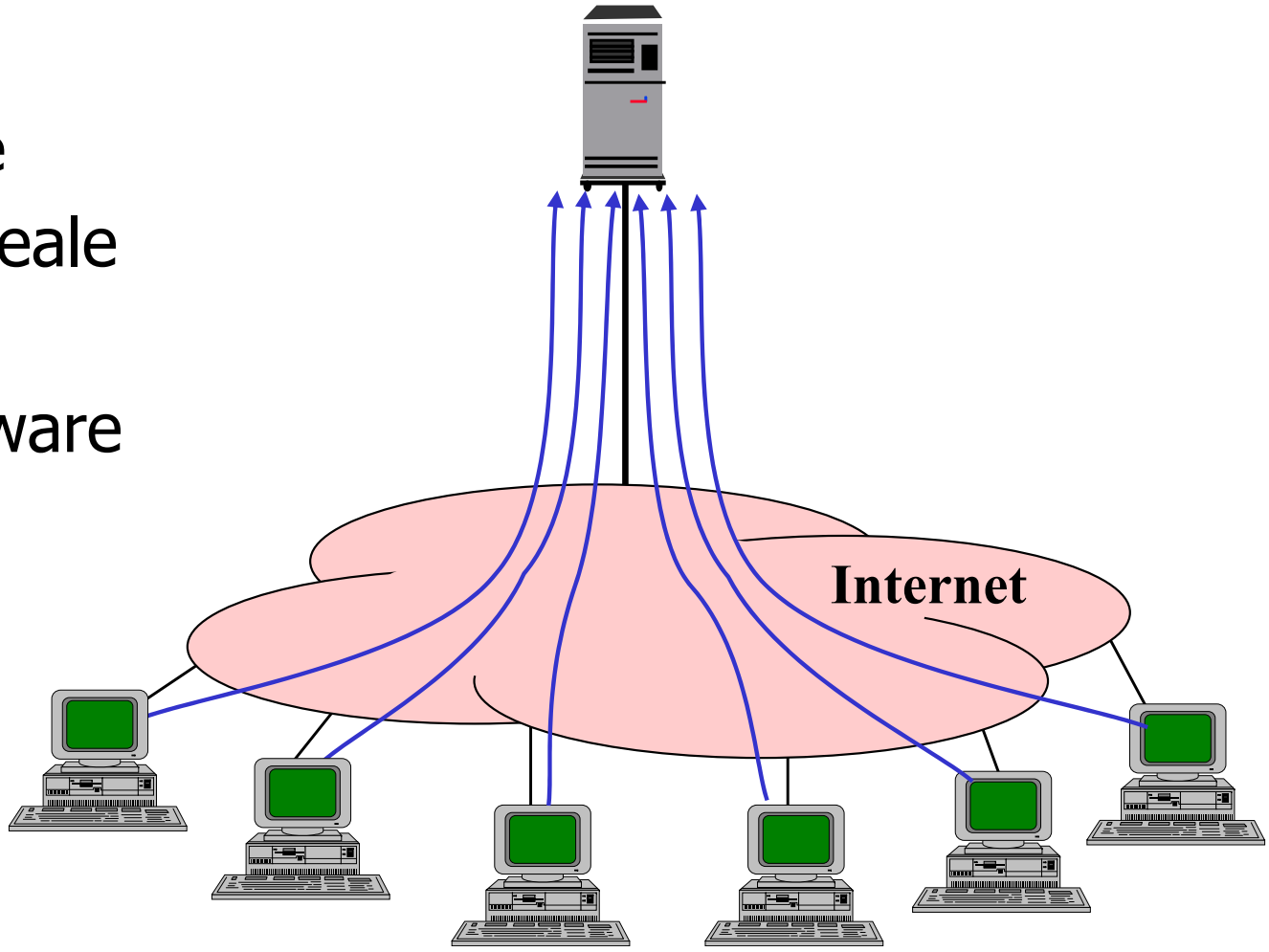
- Elaborazione hardware di informazioni di livello 4
  - In pratica porte TCP e UDP
- Spesso per *filtraggio* e/o *classificazione* (non routing)
  - Valutazione di regole per identificare pacchetti cui applicare una certa azione
    - Accodamento in una specifica coda
    - Scarto
    - Inoltro su una certa interfaccia (poco comune)
  - Le regole sono fornite in fase di configurazione

- Funzionamento wire speed su 10 Gigabit Ethernet
  - Router tradizionali realizzano funzioni di filtraggio in software con significativo abbattimento delle prestazioni
  - Per esempio, implementazione di Access Control List
- Apparati commerciali sono Multi-Layer Switch

Motivazioni per Layer 7 Switch: problemi derivanti dal successo di Internet

Moltissimi accessi contemporanei ad un server

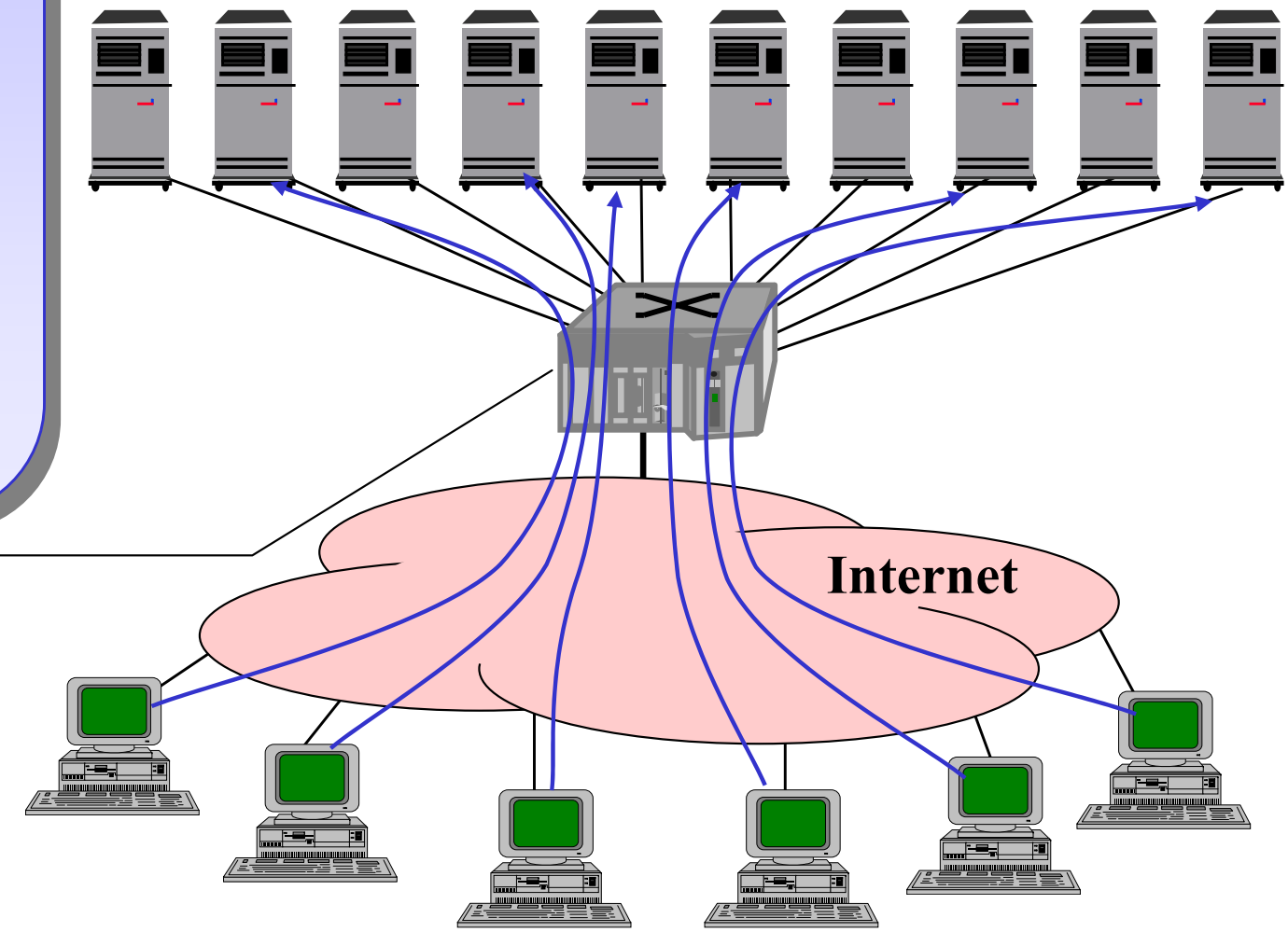
- Normalmente server web
- Portali
  - Yahoo, Netscape
- Notizie in tempo reale
  - CNN, La Stampa
- Reperimento software
  - Microsoft
- Motori di ricerca
  - Google, Yahoo



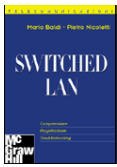
La soluzione

Server farm
Utilizzare tanti server è meno costoso che realizzare un server di sufficiente capacità

Server load balancing
Distribuendo le richieste uniformemente tra i server si crea un *server virtuale* con capacità equivalente alla somma delle capacità dei singoli server



Layer 7 switch
Application layer switch
Content aware router
Server load balancer



Application Layer Switch

- Distribuire le richieste implica un inoltro basato sul loro contenuto
  - Inoltro basato su informazioni di livello applicativo
- Necessità di operare con elevate prestazioni
  - I server della server farm hanno interfacce GE o 10GE

Hardware (ASIC) per elaborazione a livello 7

Layer 7 Switch/Application Layer Switch

Apparati commerciali sono multi-layer switch

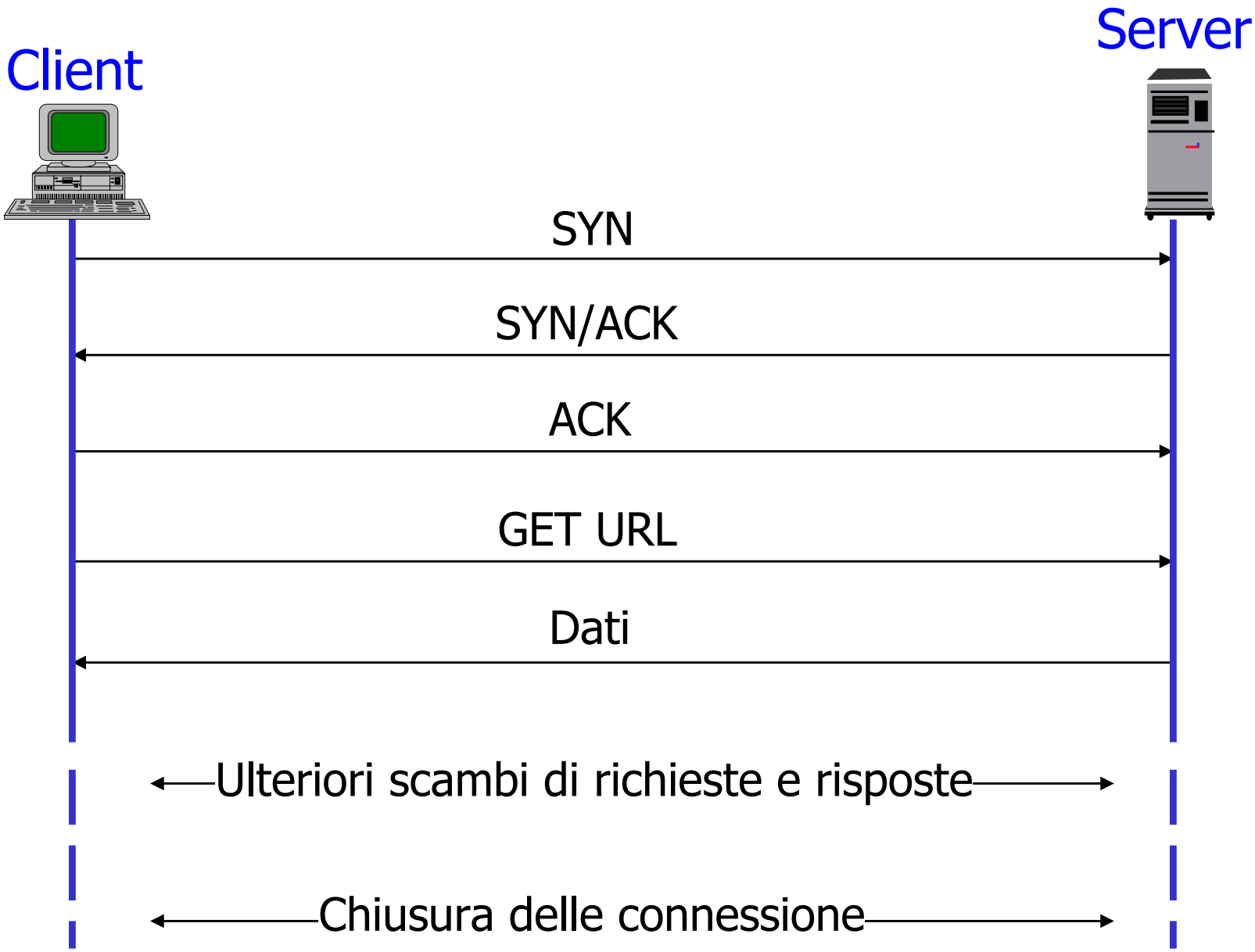
Funzionalità

- Un indirizzo IP identifica il server virtuale
- Pacchetti IP con una richiesta attivano al layer 7 switch sulla strada verso la server farm
- Layer 7 switch elabora la richiesta e decide a quale server fisico inoltrare i pacchetti IP corrispondenti
- In alcune applicazioni il server mantiene informazioni sul client → interazioni successive con lo stesso server
  - Per esempio shopping chart
  - Sticky connections (collegamenti appiccicosi)
- Limitazioni di accesso
  - Realizzazione di politiche di accesso
  - Valutazione di regole e filtraggio

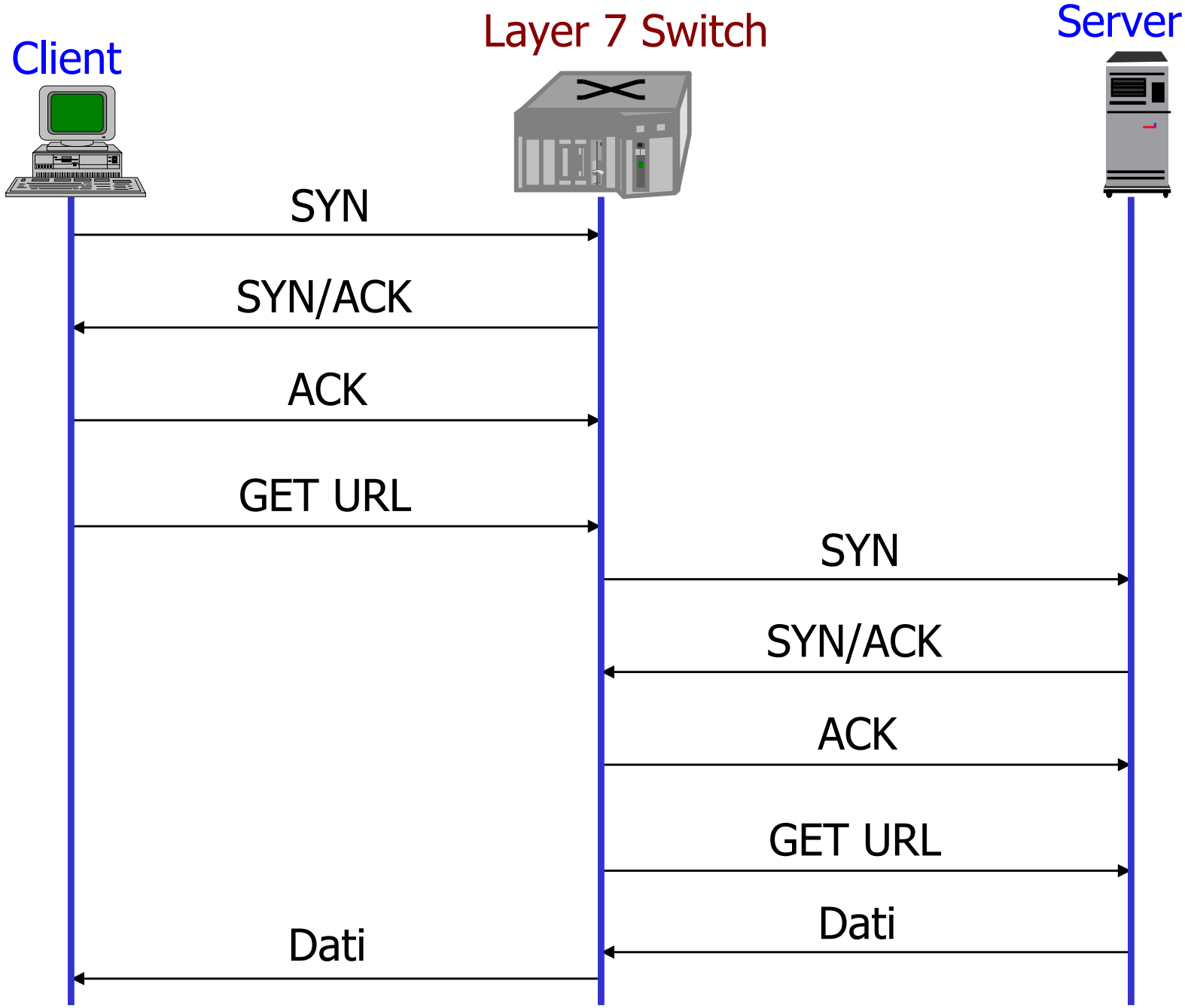
Scelta del server fisico

- Server fisici identici
  - Ognuno replica lo stesso contenuto
  - Weighted round robin
    - Nuove richieste vengono distribuite tra i server a turno
  - Weighted least connections
    - Una nuova richiesta va al server che ne sta servendo meno
  - Un peso consente di tenere conto di diverse prestazioni
- Server fisici non identici
  - Il contenuto del server virtuale è distribuito su vari server
    - Pagina di benvenuto replicata su vari server
    - Altre pagine, distribuite tra i server
  - La scelta del server dipende dal contenuto
  - Diversificazione delle richieste risulta in bilanciamento

Elaborazione di livello 7: non solamente una nuova intestazione da analizzare



Terminazione delle connessioni



Implicazioni

- Pesante lavoro di manutenzione di connessioni TCP
  - Elaborazione acknoledgements
  - Memorizzazione informazioni di stato
  - Memorizzazione pacchetti in finestra
    - Trasmessi e non ancora confermati
  - Difficile implementazione hardware

Indici di prestazioni

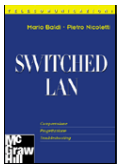
- Numero massimo di connessioni aperte
  - Qualche centinaio di migliaia
- Numero massimo di nuove connessioni al secondo
  - Qualche decina di migliaia

Scorciatoie

- Possibili quando l'inoltro delle richieste non è veramente basato sul contenuto
  - Weighted round robin, weighted least connections
  - Porta destinazione
- Individuare i messaggi di apertura di connessione
  - SYN
- Inoltrare al server reale scelto
- Apertura della connessione avviene direttamente con il server reale
- Controllare successo di apertura
- Controllare chiusura della connessione
  - Utile soprattutto per modalità weighted least connections

Funzionalità avanzate

- Controllo del corretto funzionamento dei server reali
  - Risposta appropriata a messaggi SYN
  - Raccolta ed elaborazione statistiche
- Esclusione dei server reali con malfunzionamenti
  - Controllo periodico per verificare se c'è stato recupero
- Assegnazione ad altri server di richieste pendenti su server malfunzionanti
  - Fault protection
- Controlli e statistiche sui messaggi SYN per evitare attacchi Denial of Service (DoS)
- Load balancing di layer 7 switch
  - Più apparati in parallelo
  - Coordinazione e divisione del carico di richieste



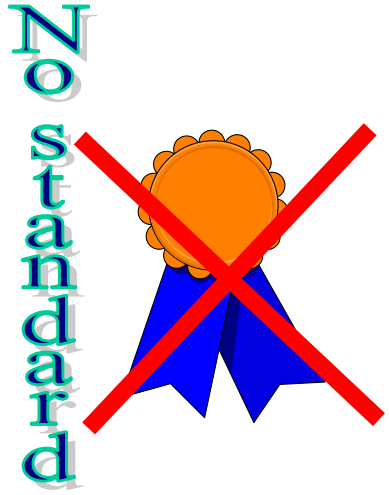
Funzionalità avanzate

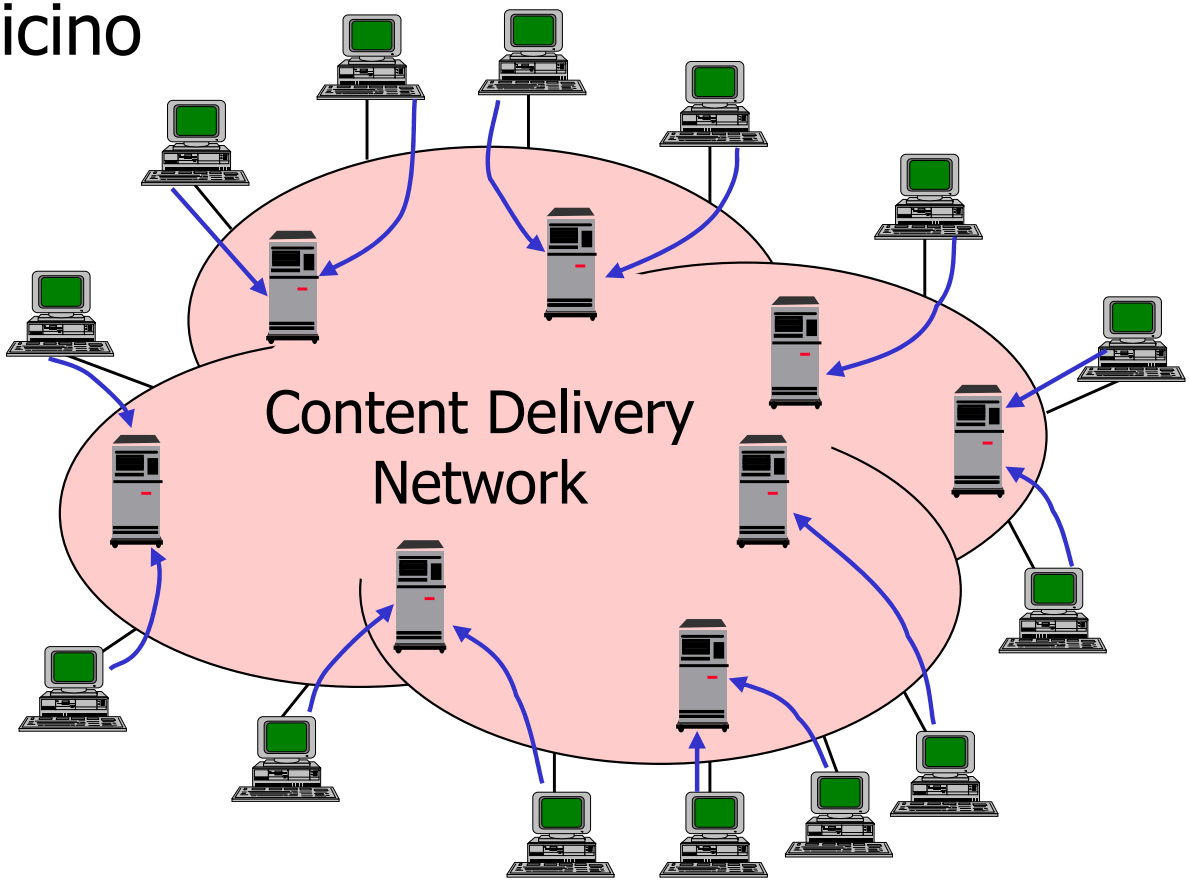
- Ridondanza di Layer 7 Switch
  - Se un apparato cessa di funzionare o resta isolato, gli altri lo sostituiscono
  - *Stateless failover*: le connessioni devono essere reinizializzate
  - *Stateful failover*: lo switch di backup conosce lo stato del primario e non deve reinizializzare le connessioni
    - Molto complicato

- Verifica dell'efficacia del bilanciamento
  - Analisi del traffico per ricavare tempi di risposta dei server
  - Un agente in esecuzione sul server raccoglie informazioni di stato e le comunica al layer 7 switch
    - Protocollo proprietario per la comunicazione con il server

Content Delivery Network (CDN)

- Generalizzazione del concetto di bilanciamento di carico
- Server farm è distribuita nella rete
- L'inoltro nella rete è basato sul contenuto delle richieste
- Si porta il contenuto vicino agli utenti
  - Riduzione dei tempi di risposta





Qualcosa di più di una diversa modalità applicazione

- In principio i Layer 7 switch sono alla base delle CDN
- Servono funzionalità avanzate di routing
  - Protocolli di routing per la disseminazione di informazioni sulla posizione di contenuti
  - Terra quasi completamente inesplorata
- Soluzioni attuali basate su espedienti
  - Minore complessità
  - Minore efficienza
- Modifiche al DNS (Domain Name Service)
  - Ad un nome corrispondono vari indirizzi
  - Fornito l'indirizzo del server più vicino al richiedente
- URL (Universal Resource Locator) adattate
  - Soluzione adottata da Akamai