Università degli Studi di Roma "Tor Vergata"
Laurea in Informatica

Sistemi Operativi e Reti
(modulo Reti)
a.a. 2023/2024

Livello di rete: piano dei dati (parte3)

dr. Manuel Fiorelli

manuel.fiorelli@uniroma2.it

https://art.uniroma2.it/fiorelli

Basate sulle slide del libro di testo:
https://gaia.cs.umass.edu/kurose_ross/ppt.php

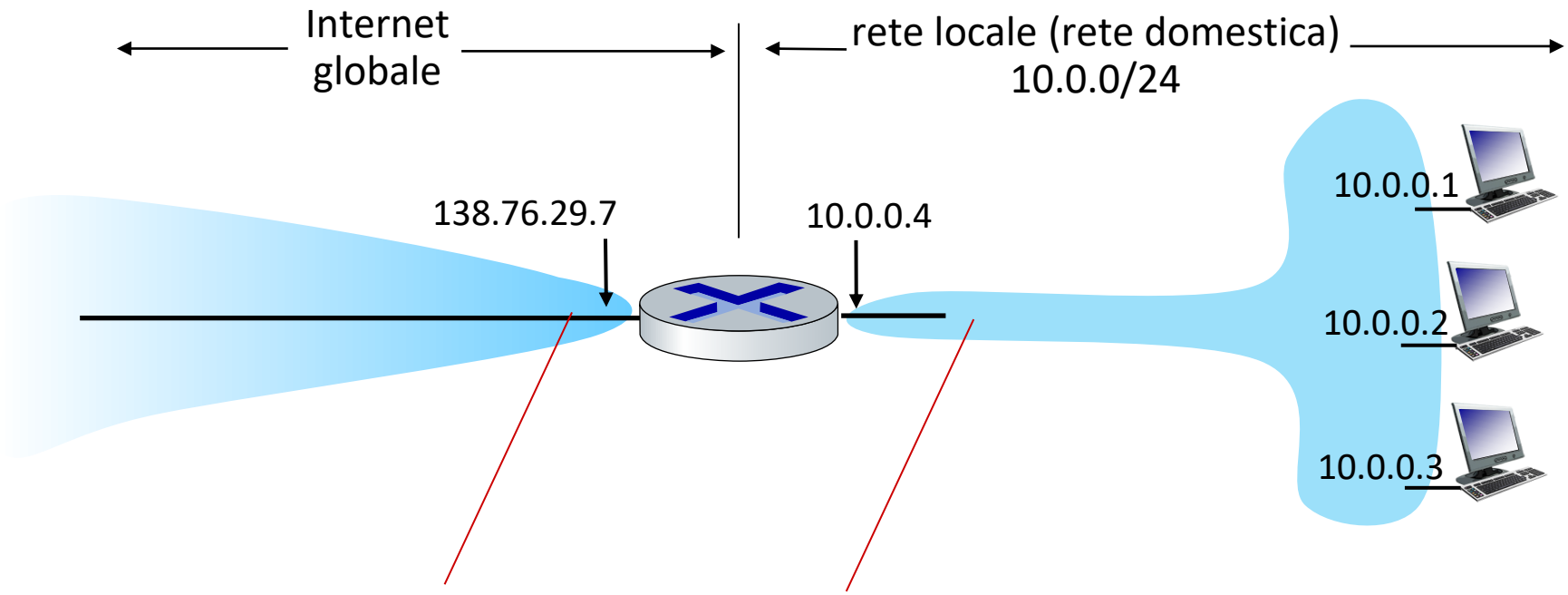
Livello di rete: tabella di marcia sul “piano dei dati”

- Livello di rete: panoramica
  - piano dei dati
  - piano di controllo
- Cosa c'è dentro un router
  - Porte di ingresso, struttura di commutazione, porte di uscita
  - buffer management, scheduling
- IP: il Protocollo Internet
  - Formato dei datagrammi
  - indirizzamento
  - Traduzione degli indirizzi di rete
  - IPv6

- Inoltro generalizzato, SDN
  - Match+action
  - OpenFlow: match+action in azione
- Middlebox

NAT: network address translation

NAT: Tutti i dispositivi della rete locale condividono un solo indirizzo IPv4 per il mondo esterno



tutti i datagrammi che *escono* dalla rete locale hanno lo *stesso* indirizzo IP sorgente: 138.76.29.7, ma *differenti* numeri di porta sorgente

I datagrammi con sorgente/destinazione in questa rete hanno indirizzo 10.0.0/24 per la sorgente e la destinazione (come al solito)

NAT: network address translation

- tutti i dispositivi della rete locale hanno indirizzi a 32 bit in uno spazio di indirizzi IP "privato" (prefissi 10/8, 172.16/12, 192.168/16) che possono essere utilizzati solo nella rete locale
- vantaggi:
  - è necessario un solo indirizzo IP dal provider ISP per *tutti* i dispositivi
  - può cambiare gli indirizzi degli host nella rete locale senza notificare il mondo esterno
  - può cambiare ISP senza modificare gli indirizzi dei dispositivi nella rete locale
  - sicurezza: dispositivi all'interno della rete locale non direttamente indirizzabili, visibili dall'esterno

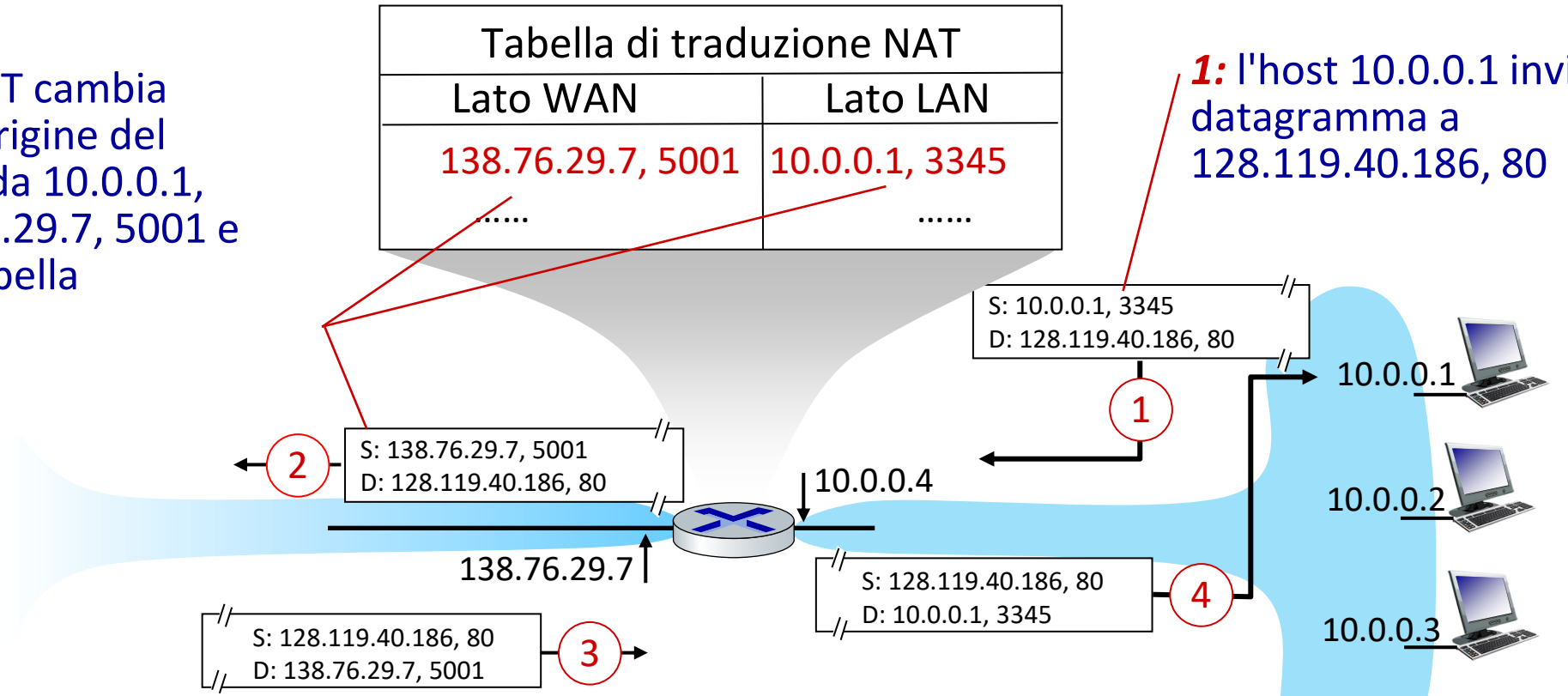
NAT: network address translation

implementazione: i router NAT devono (in maniera trasparente):

- datagrammi in uscita: sostituire (indirizzo IP sorgente, n. porta sorgente) di ogni datagramma in uscita con (indirizzo IP NAT, nuovo n. porta)
  - i client/server remoti risponderanno con (indirizzo IP NAT, nuovo n. porta) come indirizzo di destinazione
- ricordare (nella "Tabella di traduzione NAT") ogni coppia di traduzione da (indirizzo IP sorgente, n. porta) a (indirizzo IP NAT, nuovo n. porta)
- Datagrammi in ingresso: sostituire (indirizzo IP NAT, nuovo n. porta) nei campi di destinazione di ogni datagramma in ingresso con il corrispondente (indirizzo IP NAT, nuovo n. porta) memorizzato nella tabella NAT

NAT: network address translation

2: Il router NAT cambia l'indirizzo di origine del datagramma da 10.0.0.1, 3345 a 138.76.29.7, 5001 e aggiorna la tabella



3: la risposta arriva all'indirizzo di destinazione: 138.76.29.7, 5001

4: il router NAT cambia l'indirizzo di destinazione del datagramma da 138.76.29.7, 5001 a 10,0.0.1, 3345

NAT: network address translation

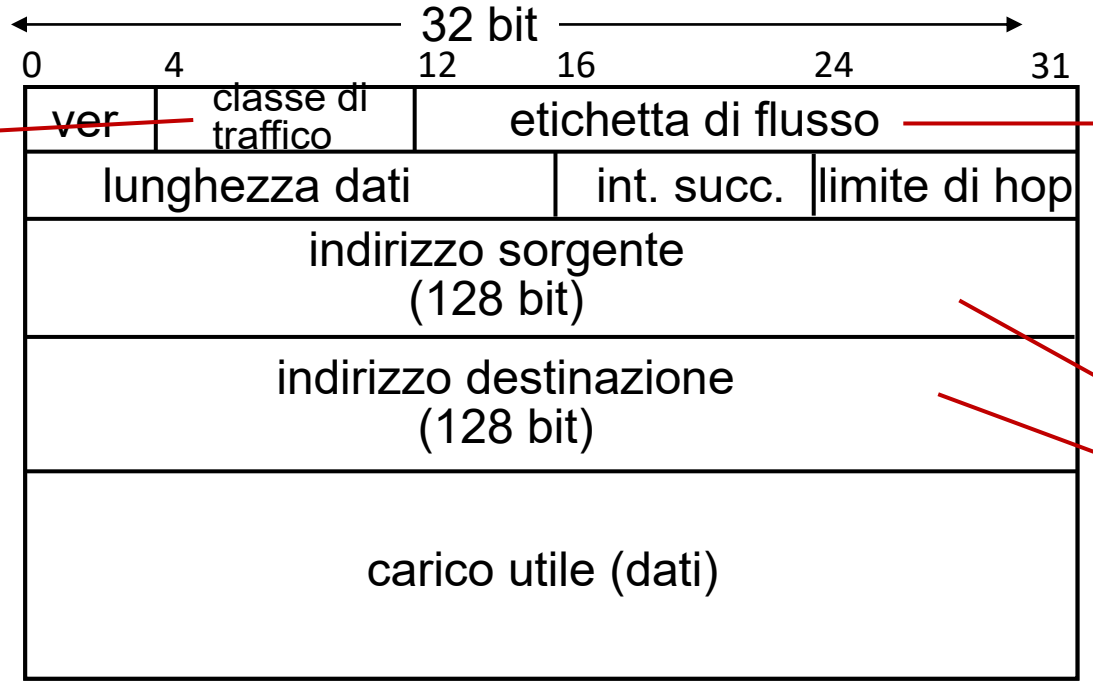
- Il NAT è oggetto di controversie:
  - i router “dovrebbero” elaborare i pacchetti solo fino al livello 3
  - la “scarsità” di indirizzi dovrebbe essere risolta da IPv6
  - viola il cosiddetto argomento punto-punto (numero di porta manipolato da un dispositivo a livello di rete)
  - attraversamento NAT (*NAT traversal*): cosa succede se un client vuole connettersi a un server dietro NAT?
- ma il NAT è qui per restare:
  - ampiamente utilizzato nelle reti domestiche e istituzionali, nelle reti cellulari 4G/5G

IPv6: motivazione

- motivazione iniziale: lo spazio degli indirizzi IPv4 a 32 bit sarebbe stato completamente allocato
- motivazioni aggiuntive:
  - velocità di elaborazione/inoltro: intestazione con una lunghezza fissa di 40 byte
  - consentire un diverso trattamento dei "flussi" a livello di rete (elevando il concetto di flusso al rango di *first-class citizen* mentre prima il focus era sui datagrammi)

Formato del datagramma IPv6

classe di traffico: attribuisce priorità a datagrammi all'interno di un flusso o proveniente da specifiche applicazioni



Etichetta di flusso: identifica i datagrammi appartenenti allo stesso flusso (concetto di "flusso" non ben definito)

128-bit indirizzi IPv6 (supporta unicast, multicast [consegna a un gruppo], anycast [consegna al più vicino di un gruppo])

Cosa manca (rispetto a IPv4):

- no checksum (per velocizzare l'elaborazione presso i router)
- no frammentazione/riassemblaggio (messaggio di errore ICMPv6 *Packet Too Big* con *MTU* del collegamento di uscita): in realtà, effettuato solo dal mittente e destinatario attraverso una *opzione*
- no opzioni (disponibile come "intestazione successiva" del protocollo di livello superiore)

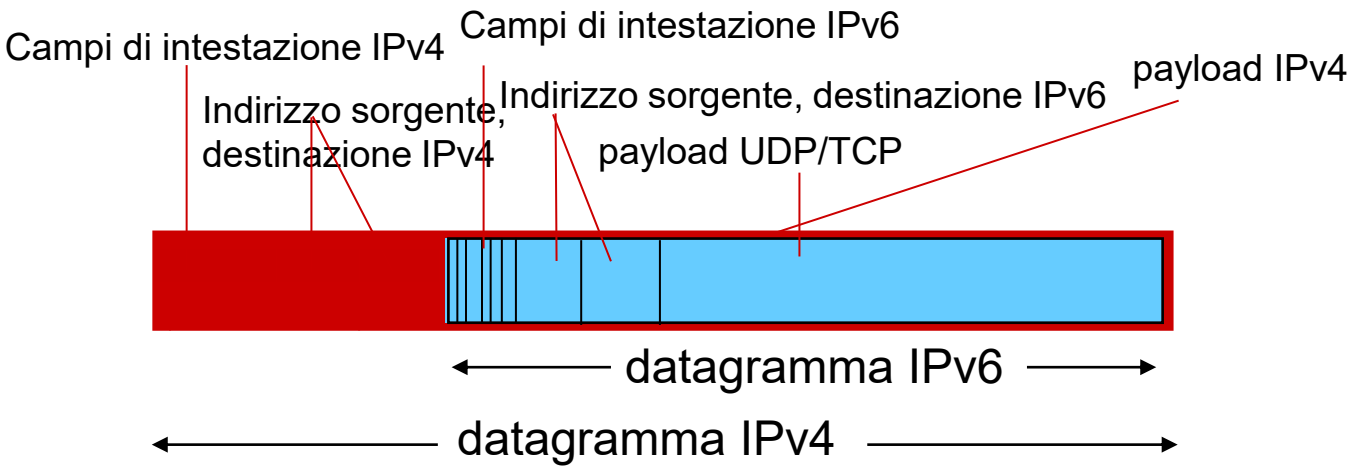
Flussi IPv6

RFC 2460 a riguardo della etichettatura dei flussi:

> l'etichettatura di pacchetti che appartengono a flussi particolari per i quali il mittente richiede una gestione speciale, come una qualità di servizio diversa da quella di default o un servizio in tempo reale"

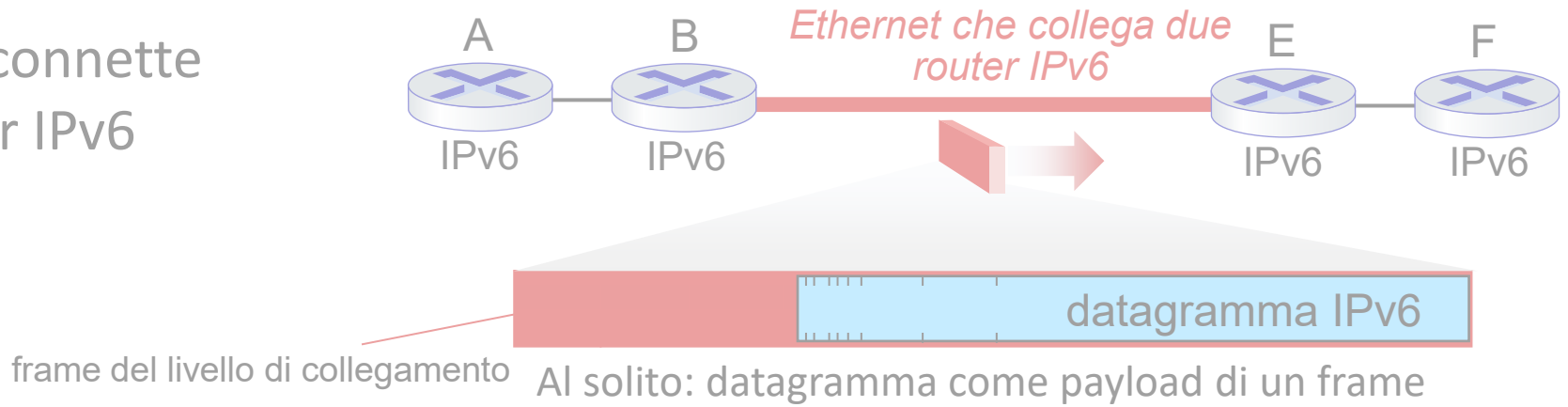
Transizione da IPv4 a IPv6

- non tutti i router possono essere aggiornati contemporaneamente
  - no “flag day” (ovvero, una "giornata campale" in cui tutte le macchine sono spente e aggiornate a IPv6)
  - come funzionerà la rete con un misto di router IPv4 e IPv6?
- tunneling: datagramma IPv6 trasportato come *payload* in un datagramma IPv4 tra i router IPv4 ("pacchetto nel pacchetto")
  - tunneling utilizzato ampiamente in altri contesti (4G/5G)

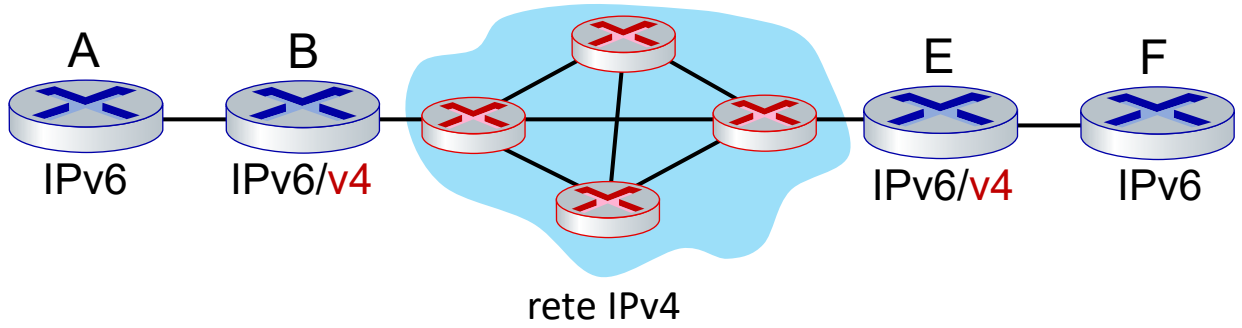


Tunneling e incapsulamento

Ethernet connette due router IPv6

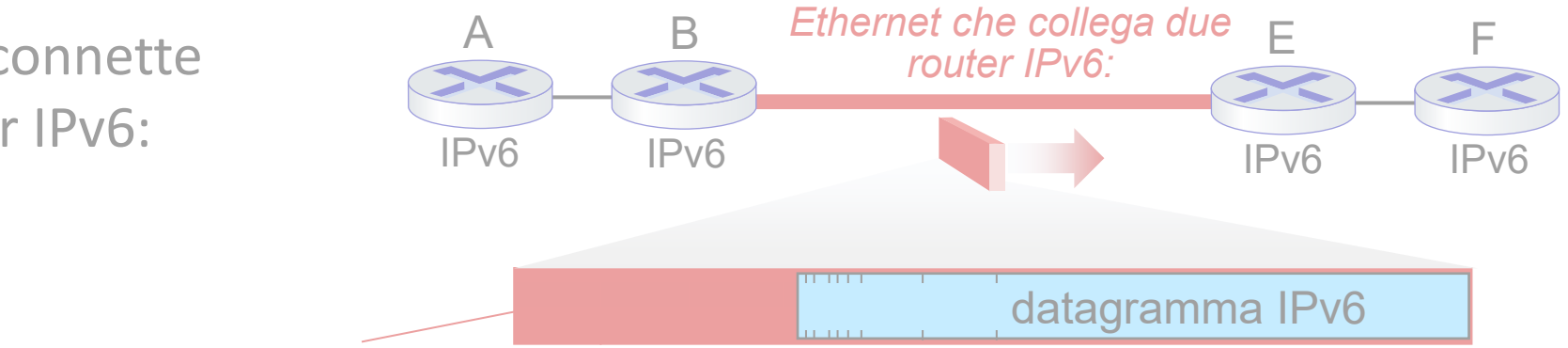


una rete IPv4 connette due router IPv6

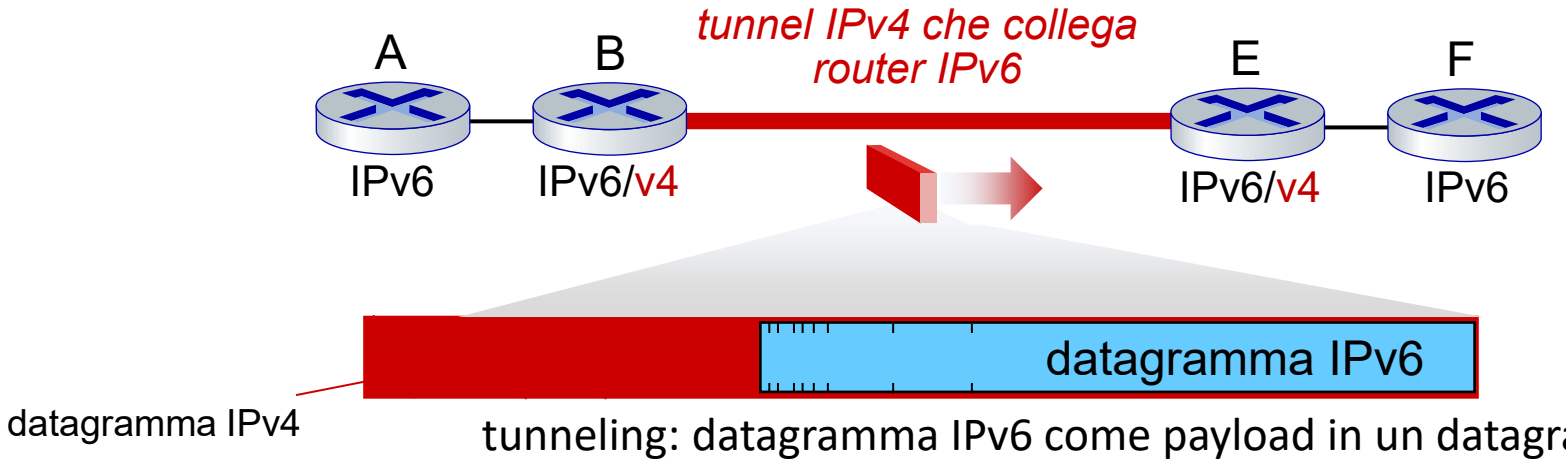


Tunneling e incapsulamento

Ethernet connette due router IPv6:

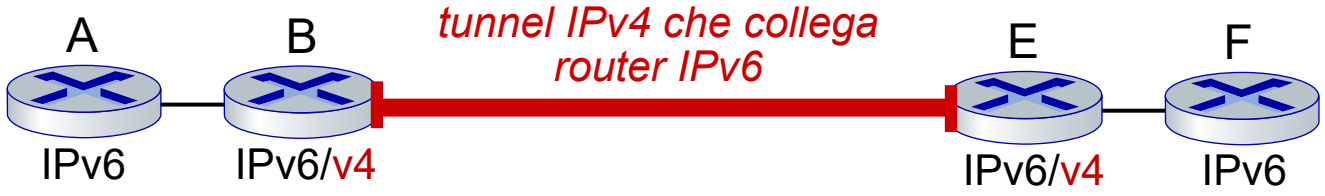


tunnel IPv4 connette due router IPv6

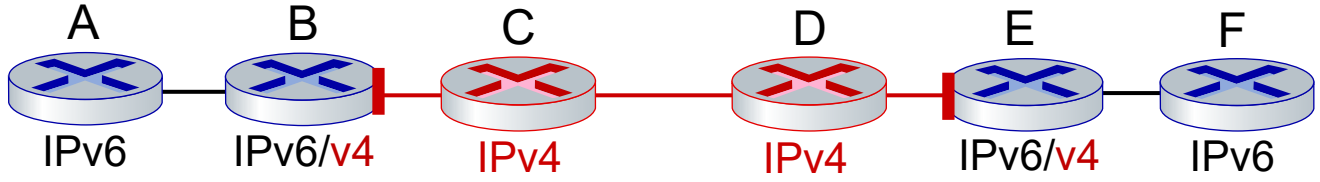


Tunneling

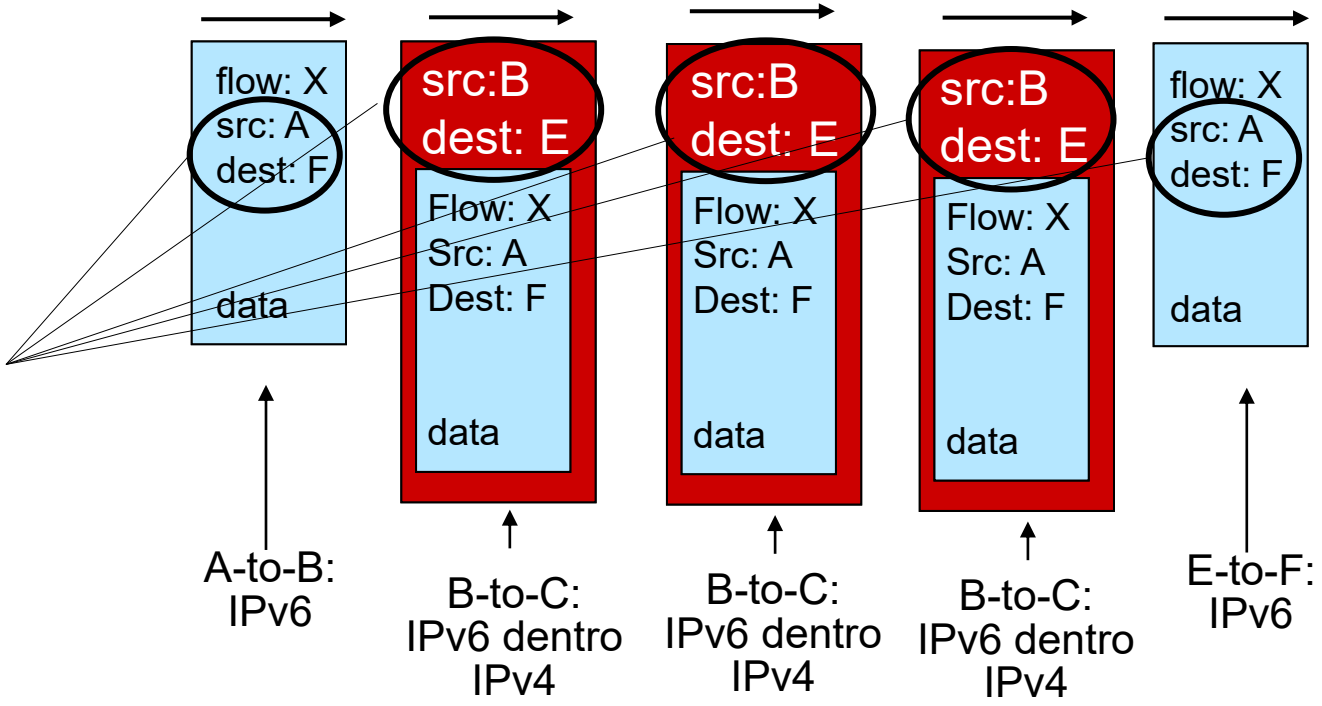
vista logica



visione fisica

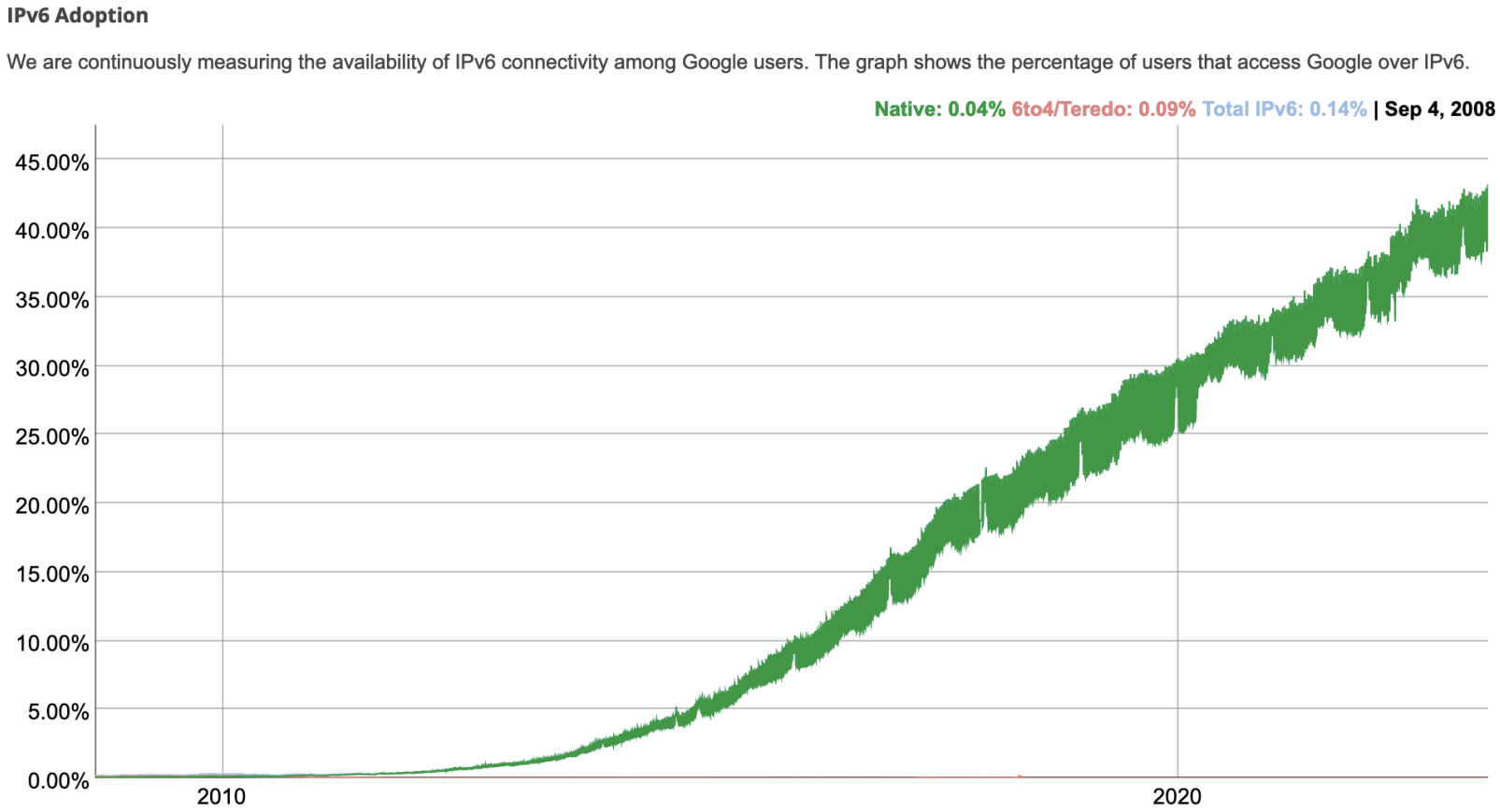


osservate gli indirizzi di destinazione



IPv6: adozione

- Google[1]: ~ 40% dei client accede ai suoi servizi attraverso IPv6 (2023)
- NIST: 1/3 di tutti i domini governativi US sono abilitati a IPv6



IPv6: adozione

- Google[1]: ~ 40% dei client accede ai suoi servizi attraverso IPv6 (2023)
- NIST: 1/3 di tutti i domini governativi US sono abilitati a IPv6
- Lungo (lunghissimo!) tempo per l'installazione e l'uso
  - 25 anni e oltre!
  - pensate ai cambiamenti a livello di applicazione negli ultimi 25 anni: WWW, social media, streaming multimediale, gaming, telepresenza, …
  - *Perché?*

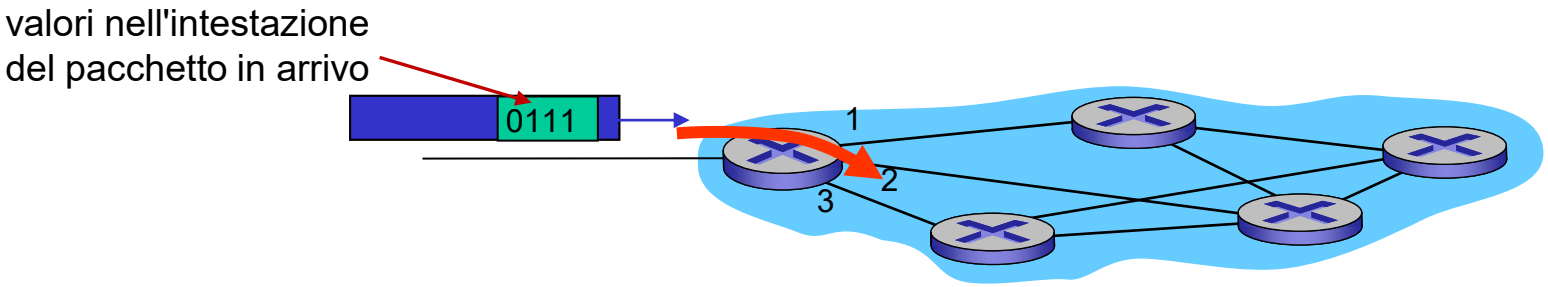
[1] https://www.google.com/intl/en/ipv6/statistics.html

Livello di rete: tabella di marcia sul “piano dei dati”

- Livello di rete: panoramica
  - piano dei dati
  - piano di controllo
- Cosa c'è dentro un router
  - Porte di ingresso, struttura di commutazione, porte di uscita
  - buffer management, scheduling
- IP: il Protocollo Internet
  - Formato dei datagrammi
  - indirizzamento
  - Traduzione degli indirizzi di rete
  - IPv6

- Inoltro generalizzato, SDN
  - Match+action
  - OpenFlow: match+action in azione
- Middlebox

Inoltro generalizzato: match plus action



Inoltro generalizzato: match plus action

Ripasso: ciascun router ha una tabella di inoltro (o: tabella dei flussi)

- astrazione "match plus action": cerca corrispondenze nei bit dei pacchetti in arrivo, agisce
  - *inoltro basato sulla destinazione:* inoltra in base all'indirizzo IP del destinatario
  - *inoltro generalizzato*:
    - più campi di intestazione posso determinare l'azione
    - più azioni possibili: scarta/copia/modifica/logga il pacchetto

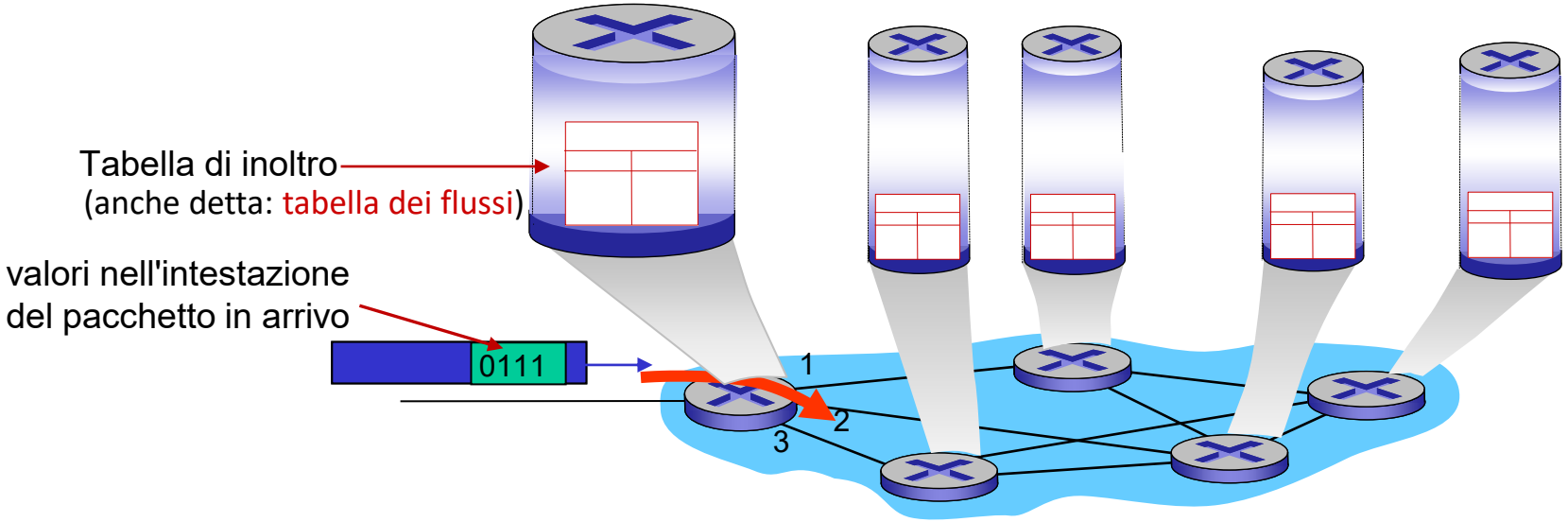
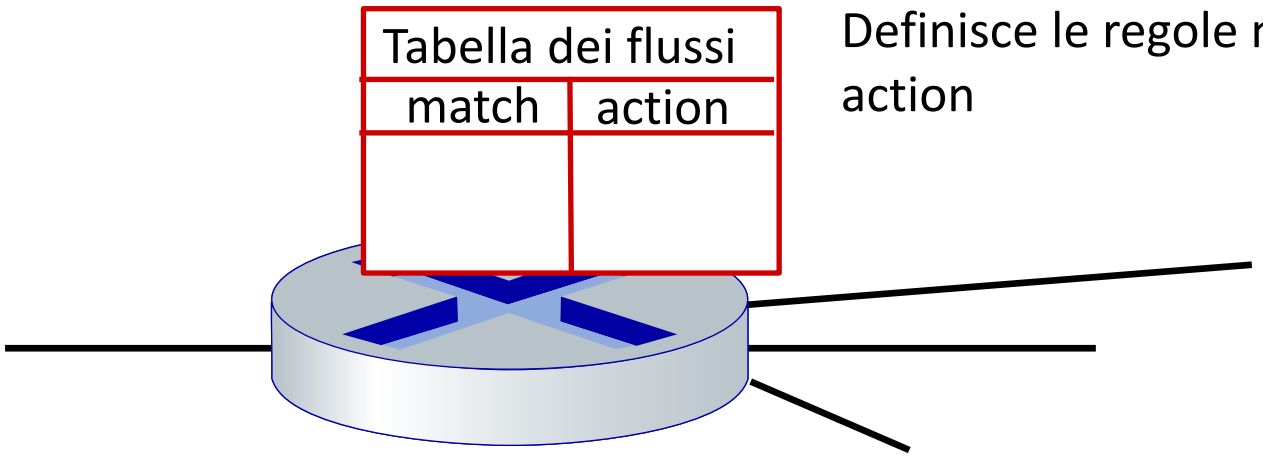


Tabella dei flussi

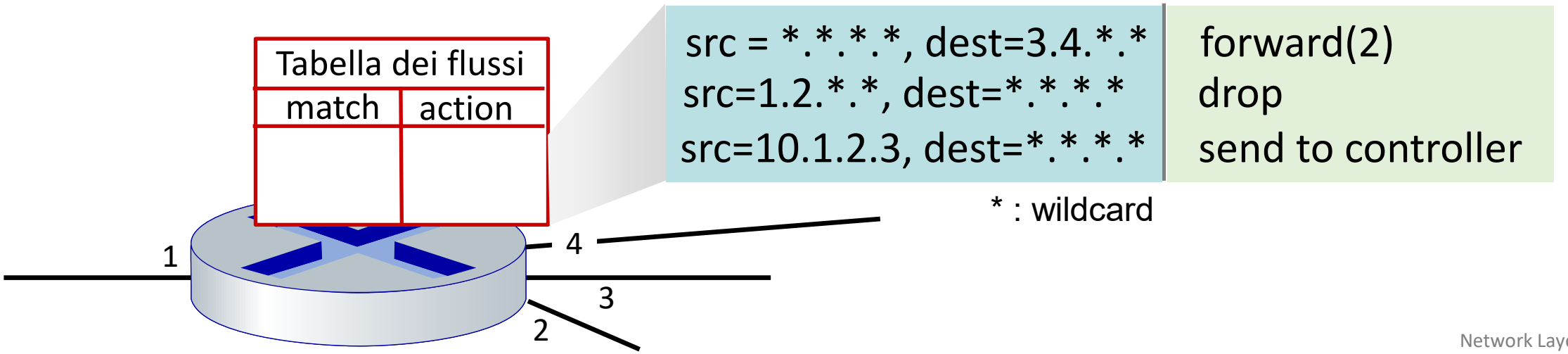
- flusso: definito dai valori campi di intestazione (a livello di collegamento, rete o trasporto)
- inoltro generalizzato: semplici regole per la gestione dei pacchetti
  - match: pattern sui valori dei campi di intestazione
  - actions: per il pacchetto in cui viene trovata una corrispondenza: scartare (drop), inoltrare (forward), modificare l'intestazione (modify), o inviare al controllore
  - priorità: disambigua pattern sovrapposti
  - contatori: numero di byte e numero di pacchetti , marca temporale ultimo aggiornamento



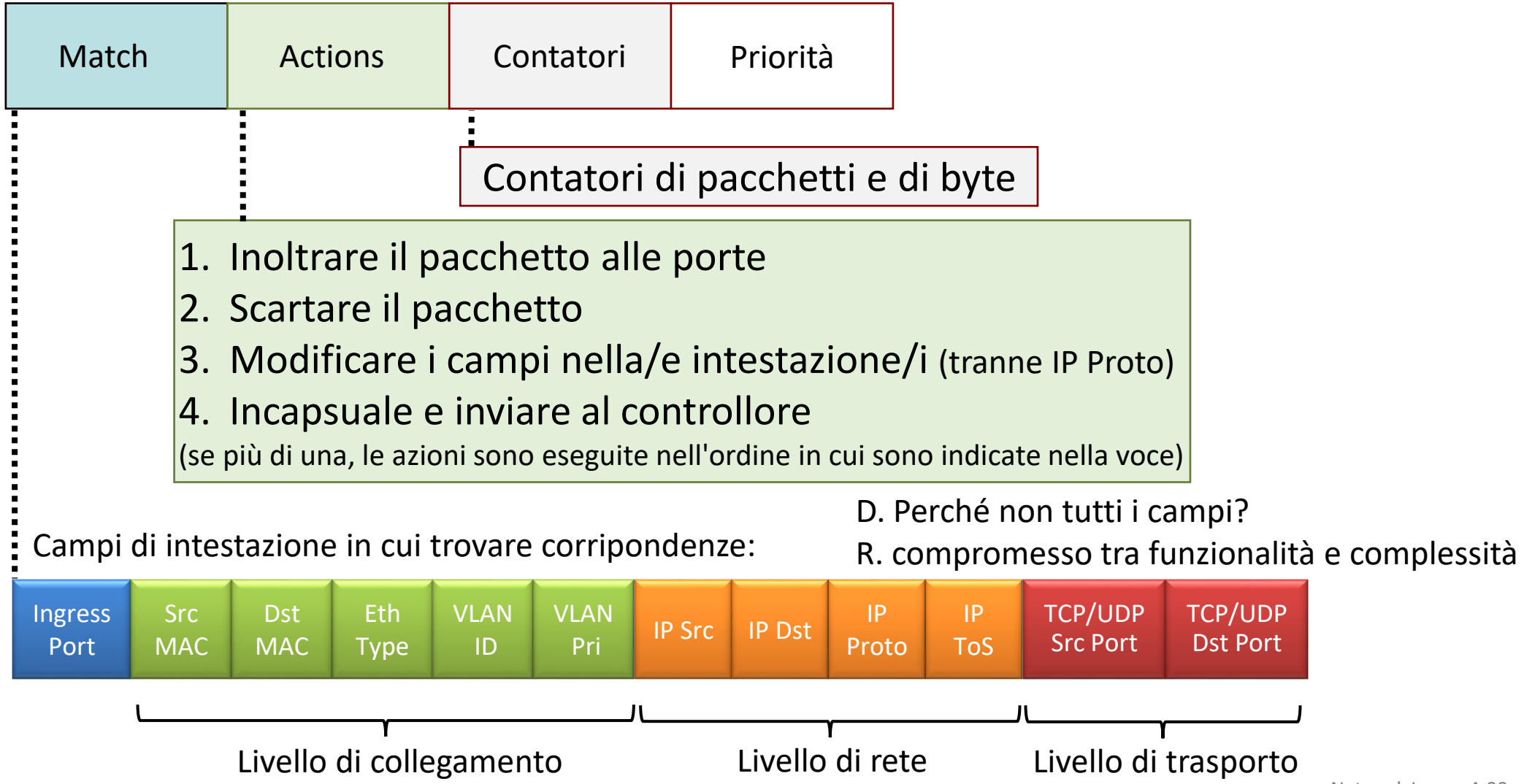
Definisce le regole match + action

Tabella dei flussi

- flusso: definito dai valori campi di intestazione (a livello di collegamento, rete o trasporto)
- inoltro generalizzato: semplici regole per la gestione dei pacchetti
  - match: pattern sui valori dei campi di intestazione
  - actions: per il pacchetto in cui viene trovata una corrispondenza: scartare (drop), inoltrare (forward), modificare l'intestazione (modify), o inviare al controllore (che può, per esempio, aggiornare la tabella dei flussi prima di restituire il pacchetto per il suo inoltro)
  - priorità: disambigua pattern sovrapposti
  - contatori: numero di byte e numero di pacchetti, marca temporale ultimo aggiornamento



OpenFlow: voci della tabella di flusso



OpenFlow: esempi

Inoltro basato sulla destinazione:

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	VLAN Pri	IP Src	IP Dst	IP Prot	IP ToS	TCP s-port	TCP d-port	Action
*	*	*	*	*	*	*	51.6.0.8	*	*	*	*	port6

I datagrammi IP destinati all'indirizzo IP 51.6.0.8 devono essere inoltrati alla porta di uscita 6 del router.

Firewall:

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	VLAN Pri	IP Src	IP Dst	IP Prot	IP ToS	TCP s-port	TCP d-port	Action
*	*	*	*	*	*	*	*	*	*	*	22	drop

Bloccare (non inoltrare) tutti i datagrammi destinati alla porta TCP 22 (numero di porta ssh)

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	VLAN Pri	IP Src	IP Dst	IP Prot	IP ToS	TCP s-port	TCP d-port	Action
*	*	*	*	*	*	128.119.1.1	*	*	*	*	*	drop

Bloccare (non inoltrare) tutti i datagrammi inviati dall'host 128.119.1.1

OpenFlow: esempi

Inoltro basato sulla destinazione a Livello 2:

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	VLAN Pri	IP Src	IP Dst	IP Prot	IP ToS	TCP s-port	TCP d-port	Action
*	*	22:A7:23: 11:E1:02	*	*	*	*	*	*	*	*	*	port3

frame di livello 2 con indirizzo MAC di destinazione 22:A7:23:11:E1:02 devono essere inoltrati alla porta di uscita 3

Load balancing

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	VLAN Pri	IP Src	IP Dst	IP Prot	IP ToS	TCP s-port	TCP d-port	Action
3	*	*	*	*	*	*	10.1.*.*	*	*	*	*	port2
4	*	*	*	*	*	*	10.1.*.*	*	*	*	*	port1

I pacchetti destinati a 10.1.*.* provenienti dalle porta 3 e 4 sono inviati rispettivamente sulle porta 2 e 1 (non possibile con l'inoltro basato sulla destinazione).

Astrazione in OpenFlow

- match+action: astrae dispositivi differenti

Router

- *match:* prefisso IP di destinazione più lungo
- *action:* inoltro (*forward*) attraverso un collegamento

Switch

- *match:* indirizzo MAC di destinazione
- *action:* inoltra (*forward*) o inonda (*flood*)

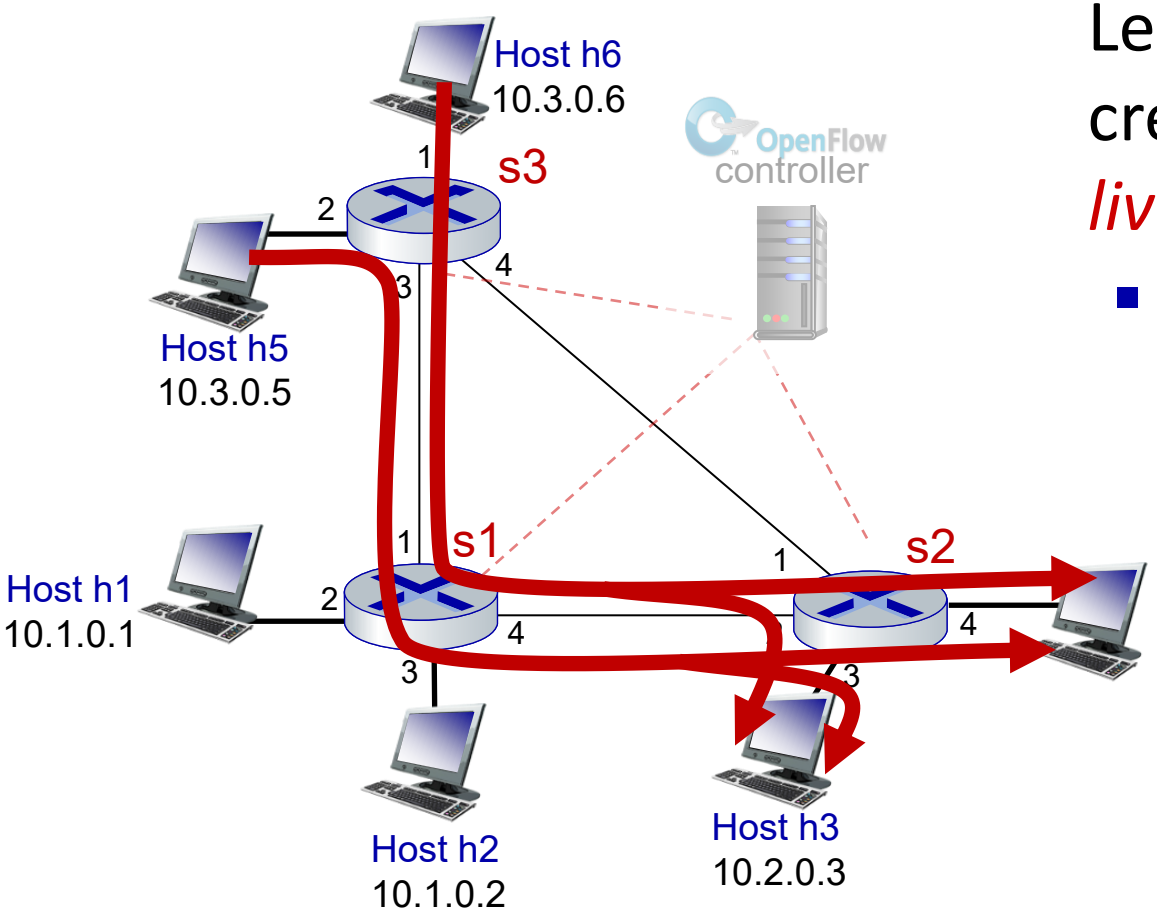
Firewall

- *match*: indirizzi IP e numeri di porta TCP/UDP
- *action:* consentire (*permit*) o negare (*deny*)

NAT

- *match:* indirizzo IP e porta
- *action:* riscrive (*rewrite*) l'indirizzo e la porta

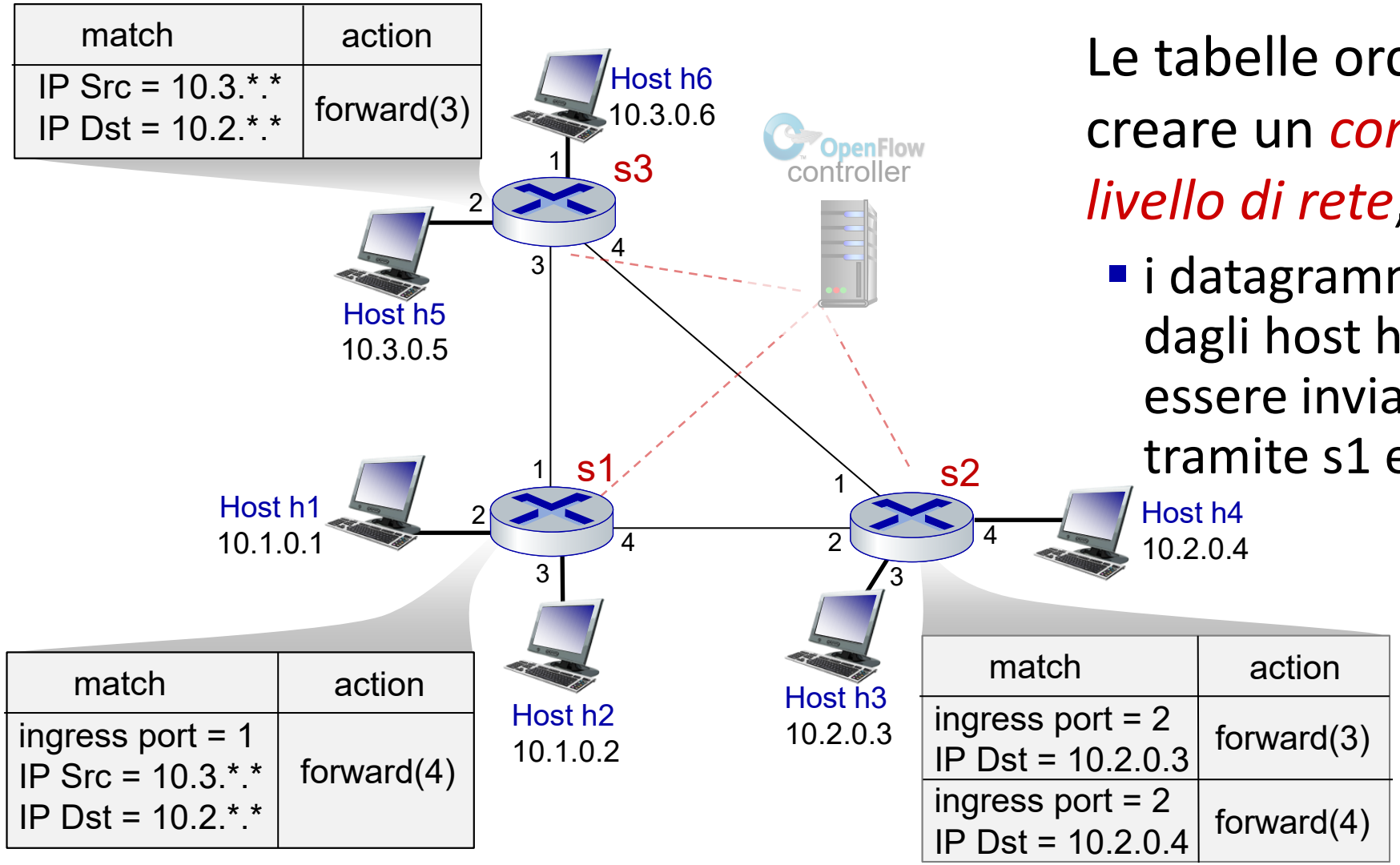
OpenFlow example



Le tabelle orchestrate possono creare un *comportamento a livello di rete*, es.,:

- i datagrammi provenienti dagli host h5 e h6 devono essere inviati a h3 o h4, tramite s1 e da qui a s2

Esempio OpenFlow



Le tabelle orchestrate possono creare un *comportamento a livello di rete*, es.,:

- i datagrammi provenienti dagli host h5 e h6 devono essere inviati a h3 o h4, tramite s1 e da qui a s2

Inoltro generalizzato: riassunto

- astrazione "match plus action": trova corrispondenze (*match*) nei bit nell'intestazione (di qualsiasi livello) dei pacchetti in arrivo, agisce (*action*)
  - trova corrispondenze su molti campi (livello di collegamento, rete, trasporto)
  - azioni locali: scarta (*drop*), inoltra (*forward*), modifica (*modify*), o invia il pacchetto al controllore
  - "programmare" *comportamenti di rete*
- una forma semplice di "programmabilità della rete"
  - "elaborazione" programmabile per pacchetto
  - *radici storiche: il networking attivo*
  - *oggi: programmazione più generalizzata*: P4 (vedi p4.org).

Livello di rete: tabella di marcia sul "piano dei dati"

- Livello di rete: panoramica
- Cosa c'è dentro un router
- IP: il Protocollo Internet
- Inoltro generalizzato
- Middlebox
  - funzioni delle middlebox
  - evoluzione e principi architetturali di Internet

Middlebox

inoltro basato sulla destinazione

Middlebox (RFC 3234)

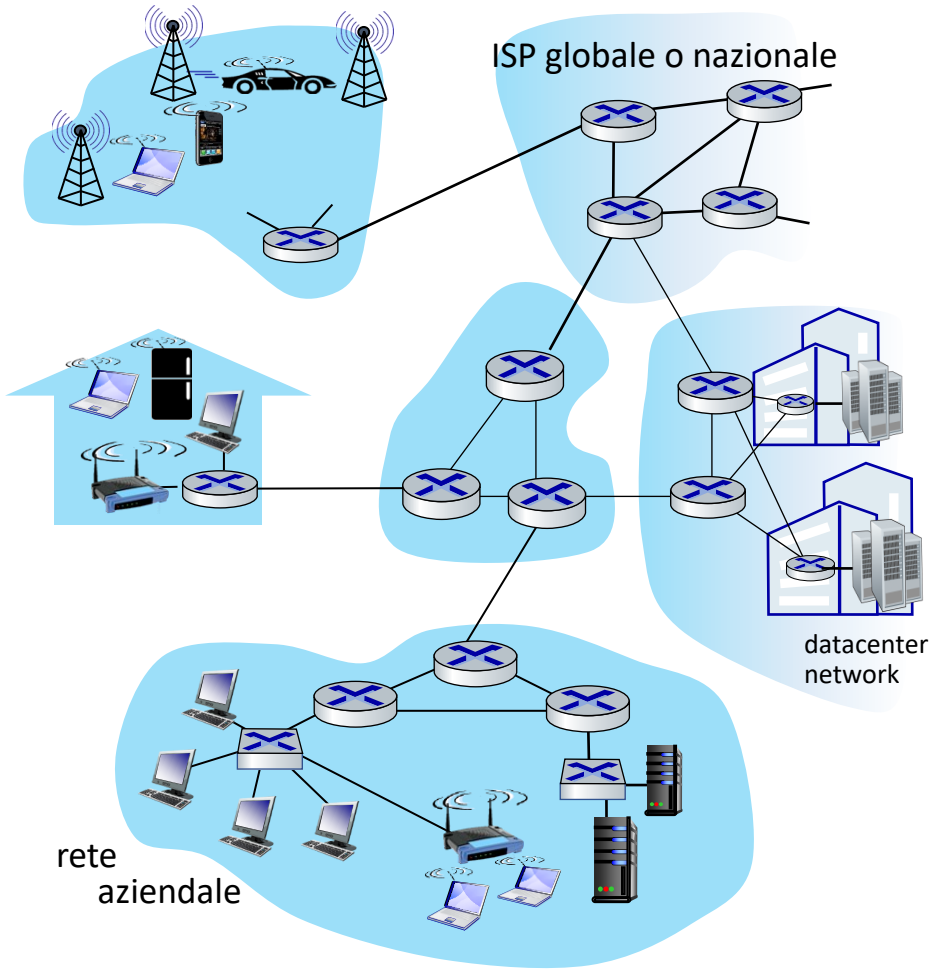
"qualsiasi box intermedio che svolge funzioni diverse da quelle normali e standard di un router IP sul percorso dei dati tra un host di origine e un host di destinazione"

si sta parlando di funzioni del piano dei dati all'interno della rete

Le middlebox sono ovunque!

NAT: nelle reti di accesso domestiche, aziendali e cellulare

Application-specific: fornitori di servizi, istituzionali, CDN



Firewalls, IDS: aziendale, istituzionale, fornitori di servizi, ISP

Load balancer: aziendale, fornitore di servizi, data center, reti mobili

Cache: fornitore di servizi, mobile, CDN

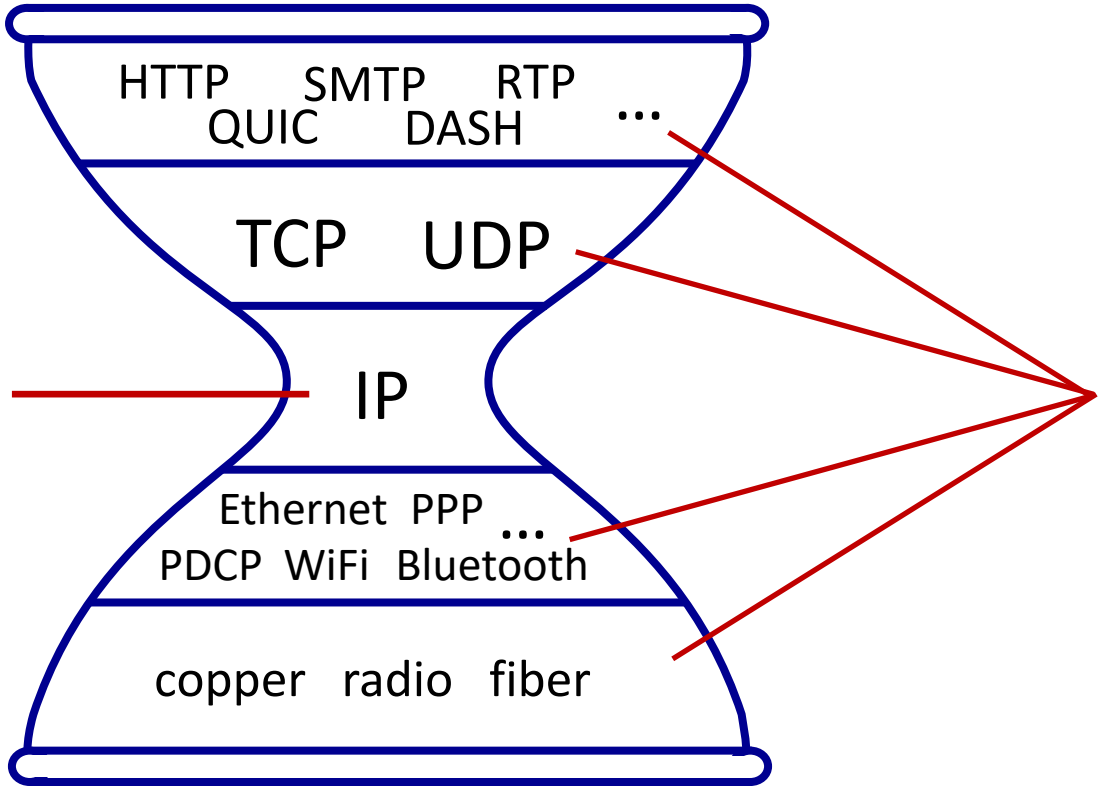
Middlebox

- inizialmente: soluzioni hardware proprietarie (chiuse)
- passaggio a hardware “whitebox” che implementa API aperte (es. OpenFlow)
  - abbandonare le soluzioni hardware proprietarie
  - azioni locali programmabili attraverso match+action
  - orientarsi verso l'innovazione/differenziazione nel software
- SDN: disaccoppia piano di controllo (centralizzato) da piano dei dati (distribuito)
- Network Functions Virtualization (NFV): astrae le funzioni di rete dall'hardware: le funzioni di rete (es. router, switch, firewall) sono programmate in software e eseguite su hardware COTS (commodity off-the-shelf) (tramite VM o container), sfruttando risorse di calcolo, storage e rete. Sono usate svariate tecniche e tecnologie per migliorare le prestazioni. Possono essere quindi anche eseguite in cloud. NFV è complementare a SDN.

Le clessidra IP

La “vita stretta” di Internet:

- *un* protocollo a livello di rete: IP
- *deve* essere implementato da ognuno dei (miliardi di) dispositivi connessi a Internet

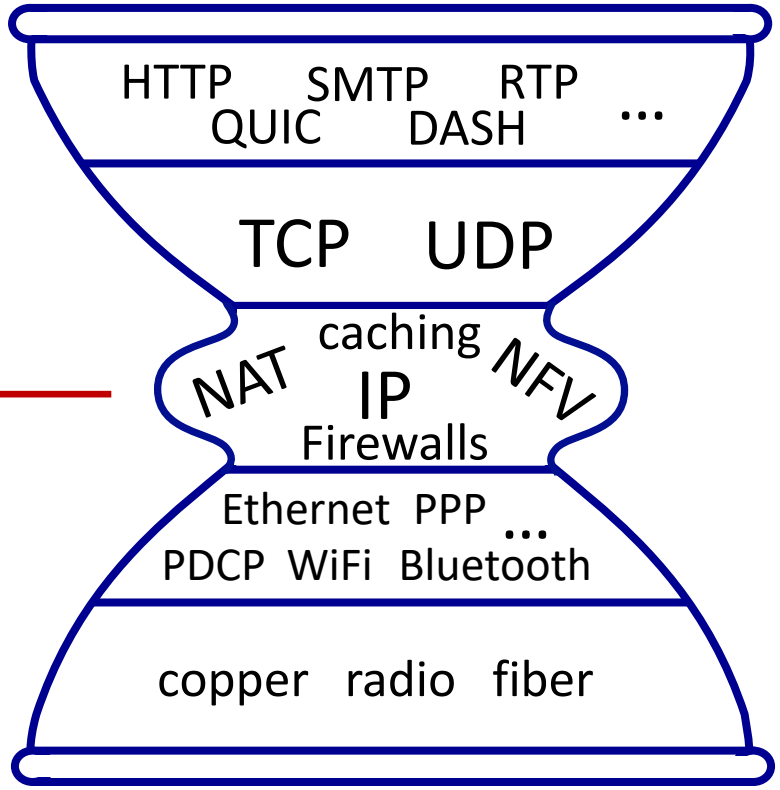


molti protocolli nei livelli di applicazione, trasporto, collegamento e fisico

La clessidra IP, alla mezza età

Le “maniglie dell'amore” della mezza età su Internet?

- middlebox, che operano all'interno della rete



Principi architetturali di Internet

RFC 1958

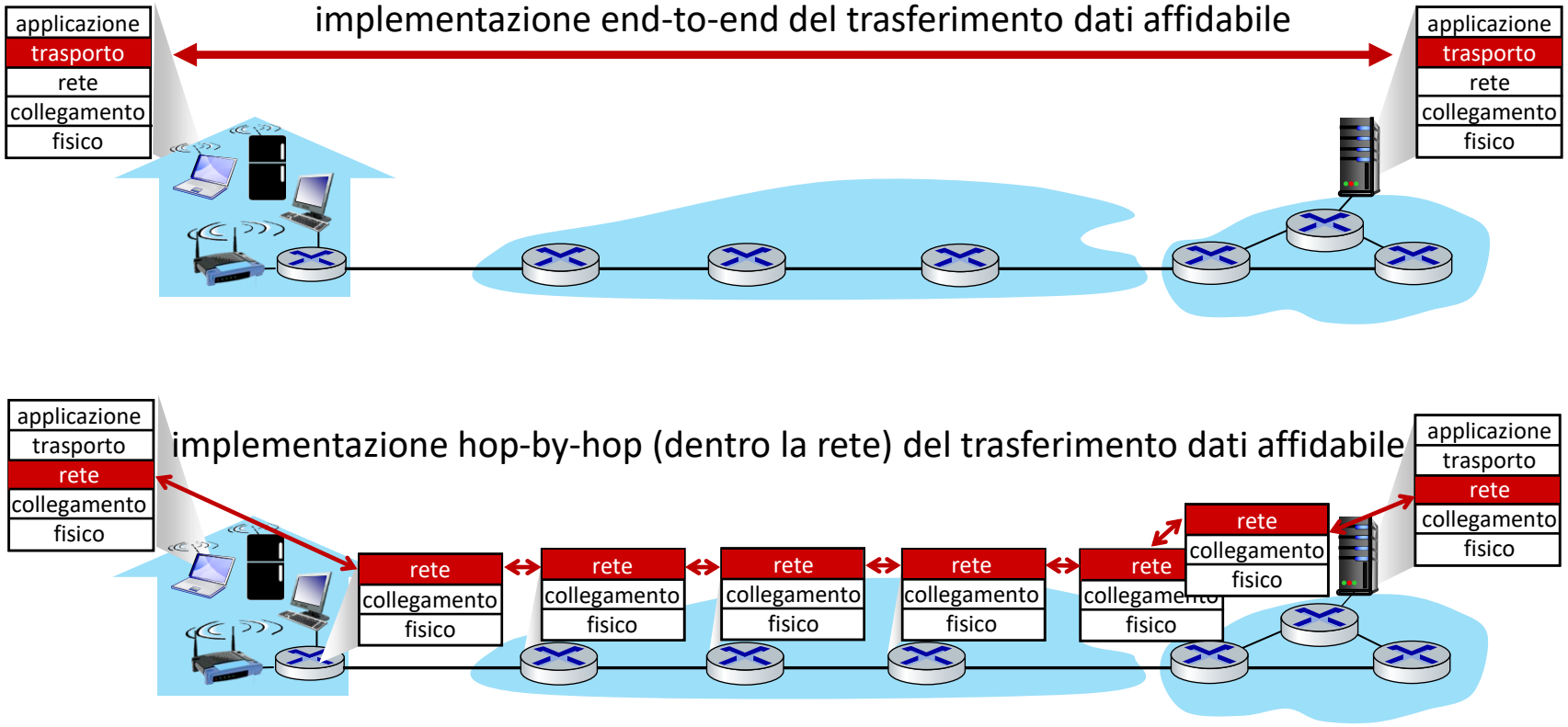
"Molti membri della comunità di Internet sostengono che non esista un'architettura, ma solo una tradizione, mai messa per iscritto per i primi 25 anni (o almeno non dallo IAB). Tuttavia, in termini molto generali, la comunità crede che l'obiettivo sia la connettività, lo strumento sia il protocollo Internet e che l'intelligenza risieda più nel paradigma end-to-end che nascosta all'interno della rete"

Tre convinzioni fondamentali:

- connettività semplice (trasferimento di datagrammi tra host)
- Protocollo IP: quella vita stretta (nasconde la eterogeneità sottostante)
- intelligenza, complessità alla periferia della rete

L'argomento "end-to-end"

- alcune funzionalità (es., trasferimento dati affidabile, controllo della congestione) possono essere implementate nel nucleo della rete o nella periferia della rete

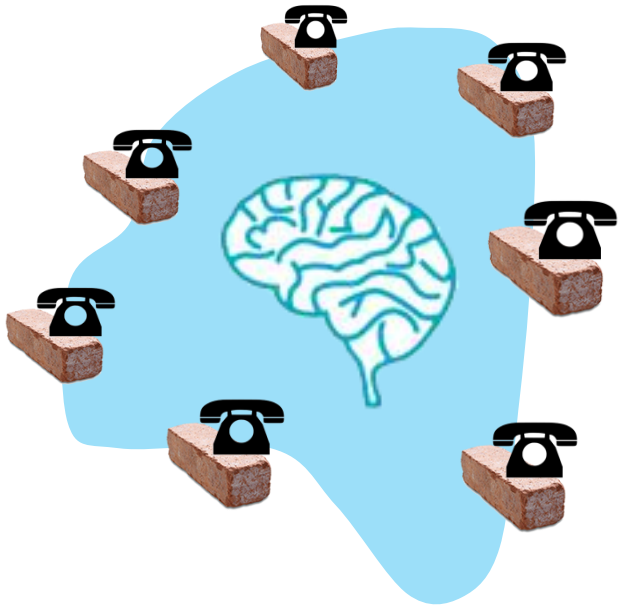


L'argomento "end-to-end"

- alcune funzionalità (es., trasferimento dati affidabile, controllo della congestione) possono essere implementate nel nucleo della rete o nella periferia della rete

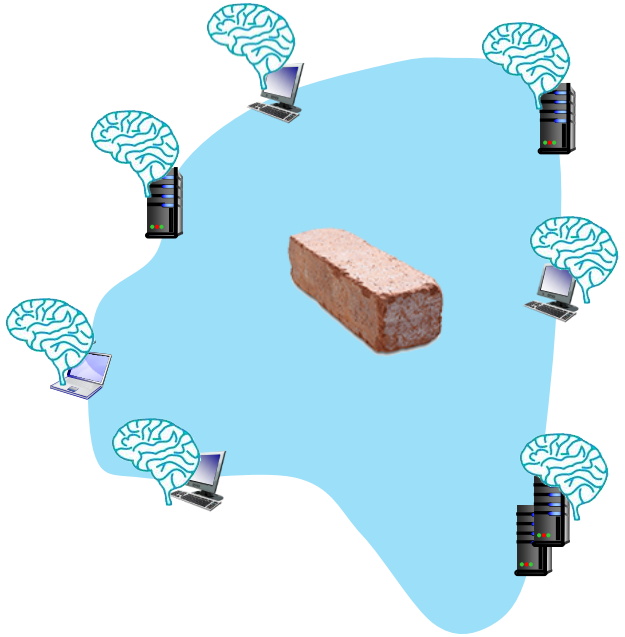
> "La funzione in questione può essere implementata completamente e correttamente solo con la conoscenza e l'aiuto dell'applicazione che sta nell'endpoint del sistema di comunicazione. Pertanto non è possibile fornire tale funzione messa in discussione come una caratteristica del sistema di comunicazione stesso (a volte una versione incompleta della funzione fornita dal sistema di comunicazione può risultare utile come potenziamento delle prestazioni).
>
> Questa linea di pensiero contraria all'implementazione delle funzioni nei livelli bassi viene chiamata "argomento end-to-end".
>
> Saltzer, Reed, Clark 1981

Dove è l'intelligenza?



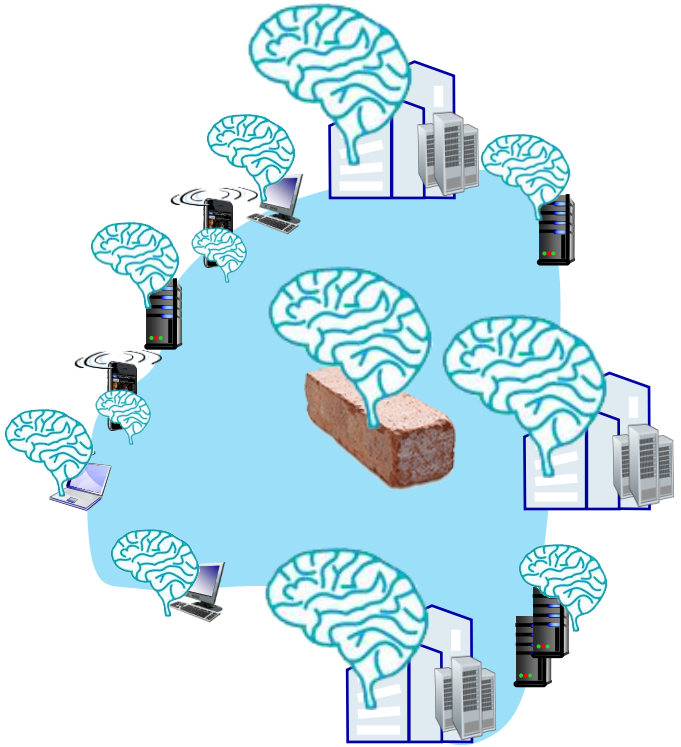
Rete telefonica del 20° secolo:

- intelligenza/calcolo negli switch di rete



Internet (pre-2005)

- intelligenza e calcolo nella periferia



Internet (post-2005)

- dispositivi di rete programmabili
- intelligenza, calcolo, infrastruttura massiccia a livello di applicazione alla periferia

Conclusione

- Livello di rete: panoramica
- Cosa c'è dentro un router
- IP: il Protocollo Internet
- Inoltro generalizzato, SDN
- Middlebox

Domanda: come sono calcolate le tabelle di inoltro (per l'inoltro basato sulla destinazione) o le tabelle dei flussi (per l'inoltro generalizzato)?

Risposta: dal piano di controllo