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

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

Livello di rete: piano di controllo (parte2)

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 del “piano di controllo”

- introduzione
- algoritmi di instradamento
  - link state
  - distance vector
- instradamento interno al sistema autonomo: OSPF
- instradamento tra sistemi autonomi: BGP
- piano di controllo SDN
- Internet Control Message Protocol

- gestione e configurazione della rete
  - SNMP
  - NETCONF/YANG

Rendere l'instradamento scalabile

il nostro studio di routing fino ad ora - idealizzato

- tutti i router sono identici
- la rete è “piatta”

… non è vero nella pratica

scalabilità: miliardi di destinazioni:

- non può memorizzare tutte le destinazioni nelle tabelle di routing!
- lo scambio di tabelle di instradamento ingolferebbe i collegamenti!
- gli algoritmi distance vector impiegherebbe un tempo enorme per convergere!

autonomia amministrativa:

- Internet: una rete di reti
- ogni amministratore di rete può voler controllare l'instradamento nella propria rete o nascondere dettagli della sua struttura interna

Approccio di Internet al routing scalabile

aggregare i router in regioni note come "sistemi autonomi" (AS, *autonomous system*) (anche detti "domini"): di solito formati da router sotti alla stessa amministrazione.

Un ISP può cosituire un unico AS oppure essere partizionato in più AS.

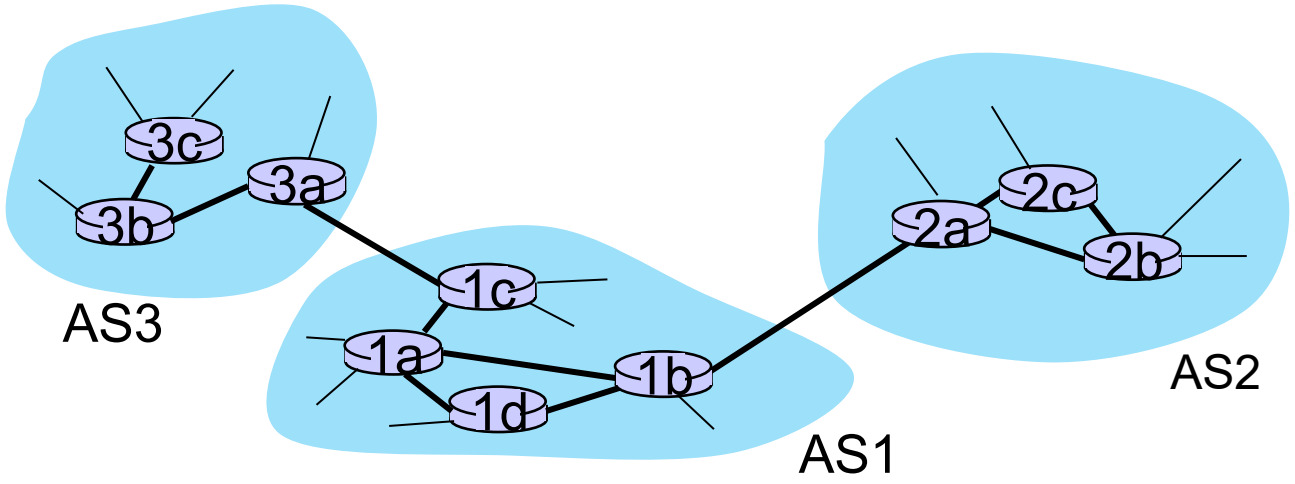
intra-AS (o "intra-domain"):
instradamento *interno al sistema autonomo ("rete")*

- tutti i router nell'AS devono eseguire lo stesso *protocollo di instradamento interno al sistema autonomo*
- tutti in AS differenti possono eseguire differenti *protocolli di instradamento interno al sistema autonomo*
- router gateway: sul "bordo" (*edge*) del proprio AS, connesso a uno o pià router in altri AS

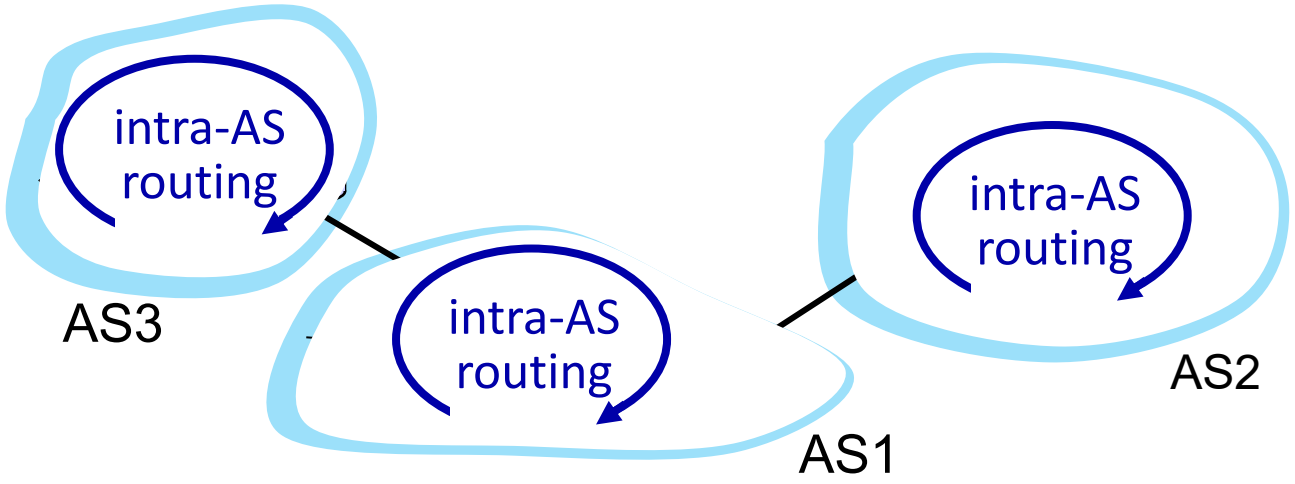
inter-AS (o "inter-domain"):
instradamento *tra* sistemi autonomi

- i gateway effettuano l'instradamento inter-AS (come pure l'instradamento intra-AS)

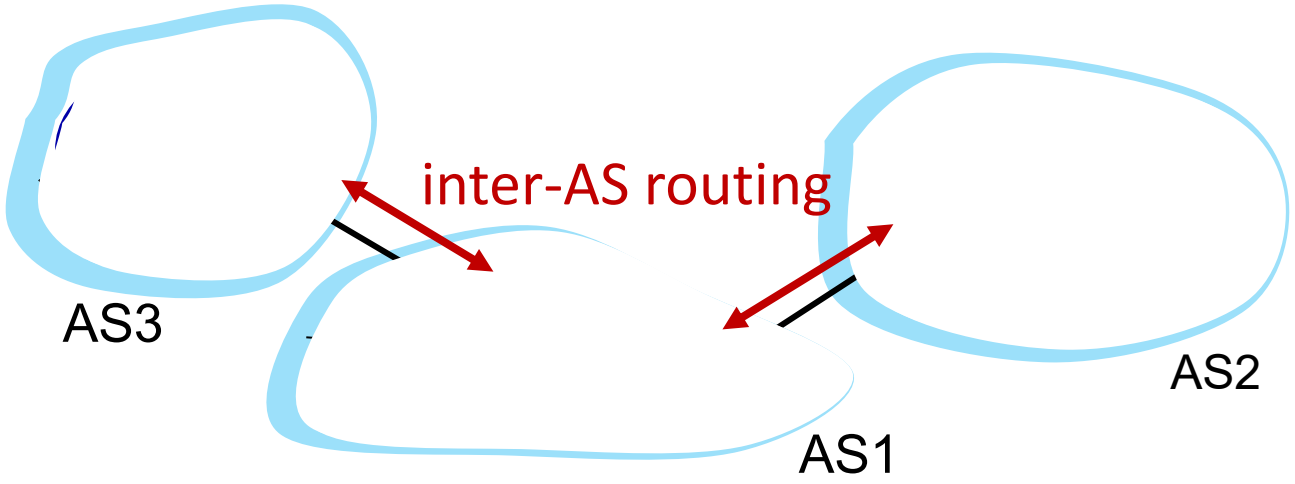
Sistemi autonomi (AS) interconnessi



Sistemi autonomi (AS) interconnessi



Sistemi autonomi (AS) interconnessi



Sistemi autonomi (AS) interconnessi

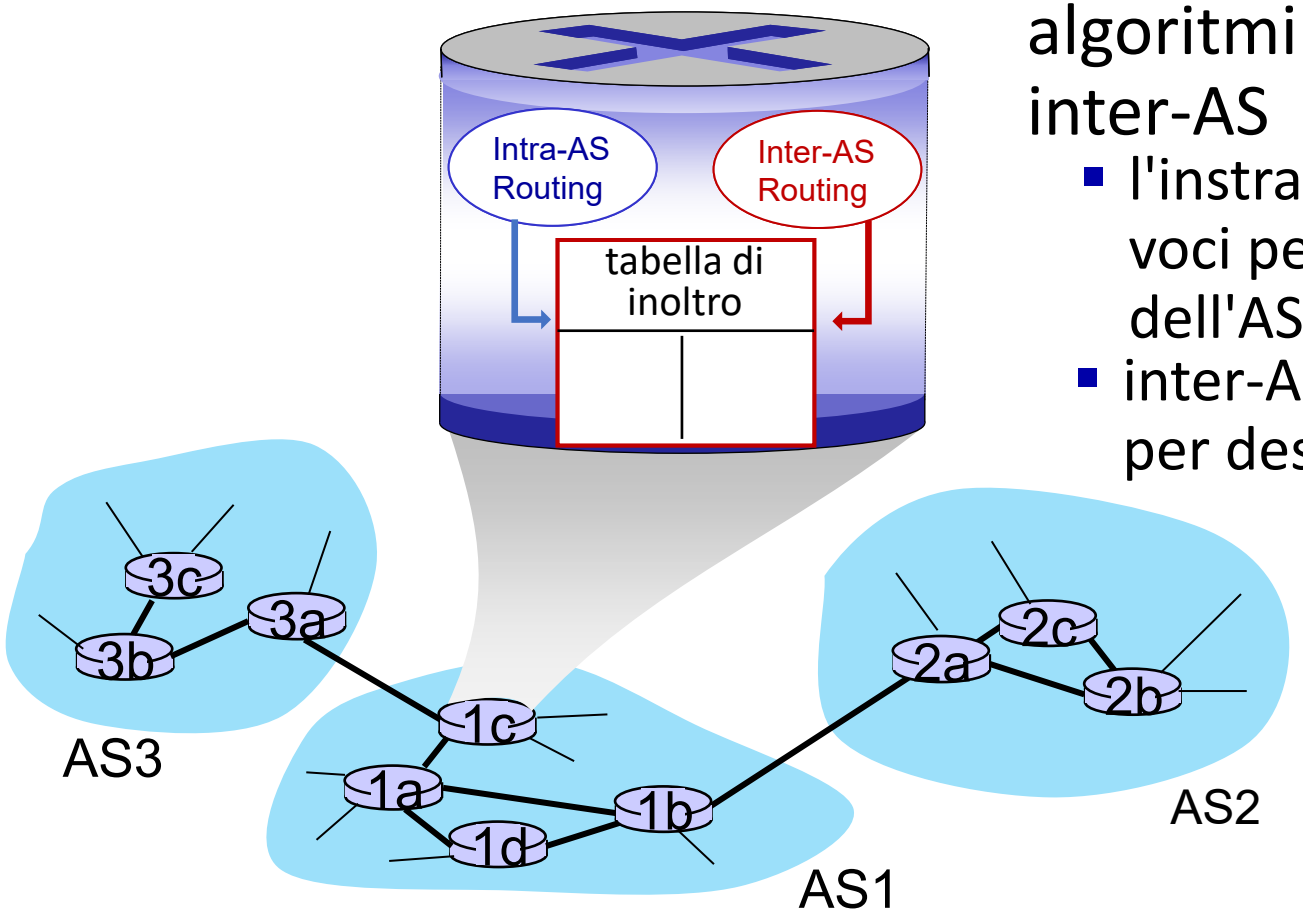


tabella di inoltro configurata dagli algoritmi di instradamento intra- e inter-AS

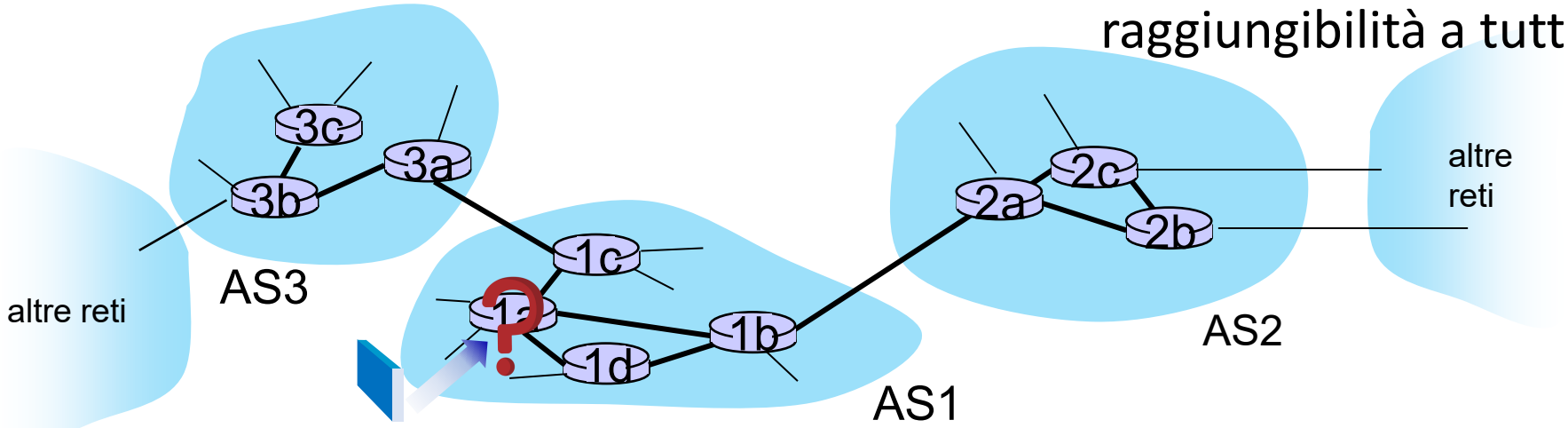
- l'instradamento intra-AS determina le voci per le destinazioni all'interno dell'AS
- inter-AS & intra-AS determinano le voci per destinazioni esterne

Instradamento inter-AS: un ruolo nell'inoltro intra-dominio

- Si supponga che un router dentro AS1 riceva un datagramma destinato al di fuori di AS1:
  - Il router dovrebbe inoltrare il pacchetto a un router gateway in AS1, ma quale?

l'instradamento inter-AS in AS1 deve:

1. imparare quali destinazioni sono raggiungibili attraverso AS2 e quali attraverso AS3
2. propagare queste informazioni di raggiungibilità a tutti i router in AS1



Instradamento intra-AS: instradamento interno al AS

protocolli di instradamento intra-AS più comuni:

- RIP: Routing Information Protocol [RFC 1723]
  - DV classico: DV scambiati ogni 30 secondi
  - non più largamente usato
- EIGRP: Enhanced Interior Gateway Routing Protocol
  - basato su DV
  - precedentemente di proprietà di Cisco per decenni (è diventato aperto nel 2013 [RFC 7868])
- OSPF: Open Shortest Path First [RFC 2328]
  - instradamento link-state
  - Protocollo IS-IS (ISO standard, non standard RFC) essenzialmente identico a OSPF

OSPF (Open Shortest Path First)

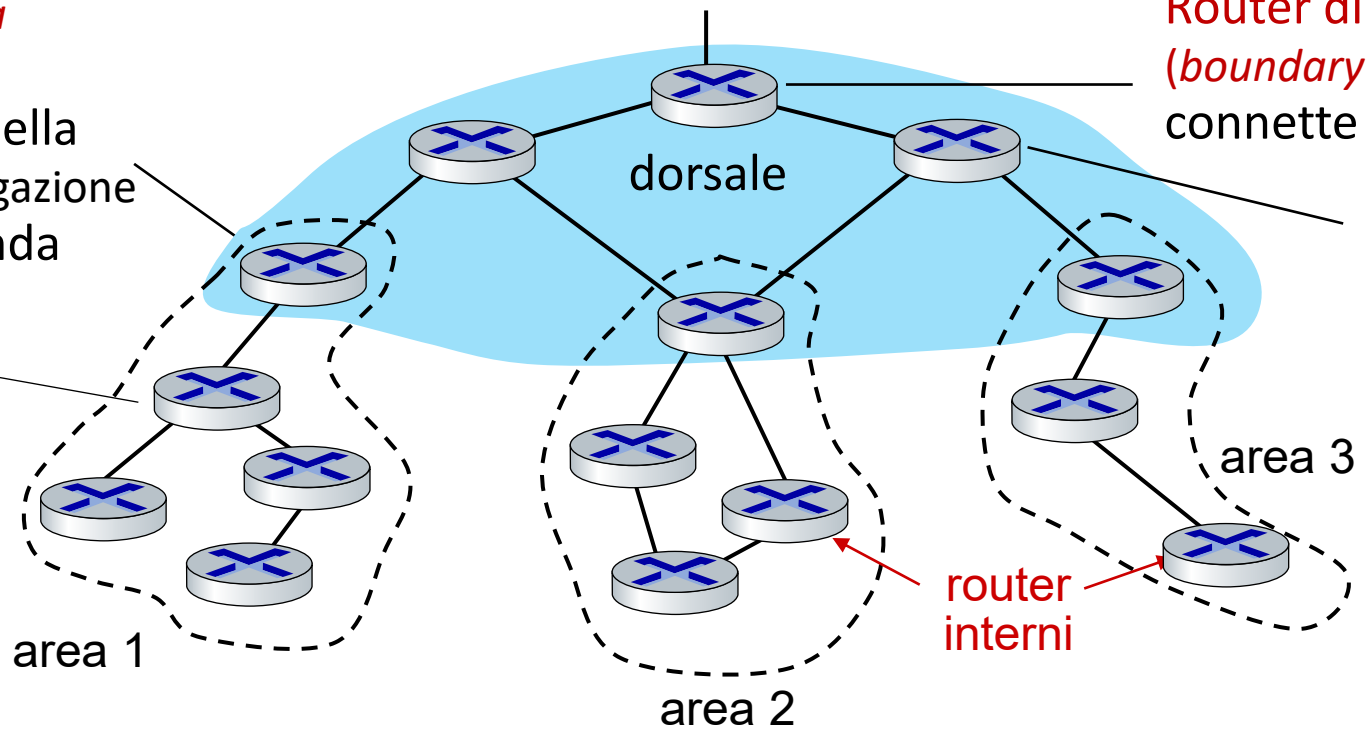
- “aperto”: disponibile pubblicamente
- classico link-state
  - ciascun router utilizza il flooding (inondazione) per inviare in broadcast le informazioni circa lo stato dei collegamenti (direttamente su IP invece di utilizzare TCP/UDP) a tutti gli altri router nell'intero AS
  - è possibile utilizzare più metriche di costo del collegamento: larghezza di banda, ritardo
  - ogni router dispone di una topologia completa, utilizza l'algoritmo di Dijkstra per calcolare la tabella di inoltro
- *sicurezza:* tutti i messaggi OSPF sono autenticati (per prevenire intrusioni dannose)

OSPF gerarchico

- gerarchia a due livelli: area locale, dorsale (*backbone*).
  - annunci link-state inondati solo in area o dorsale
  - ogni nodo ha una topologia dettagliata dell'area; conosce solo la direzione per raggiungere altre destinazioni

router di confine d'area (*area border routers*): "riassume" le distanze verso destinazioni nella propria area (in realtà, l'aggregazione degli indirizzi è opzionale), manda annunci nella dorsale

router locali (*local routers*):
- inviano le informazioni di stato solo nell'area
- calcolano l'instradamento dentro l'area
- inoltrano i pacchetti all'esterno attraverso i router di confine d'area



Router di confine (*boundary router*): si connette ad altri AS

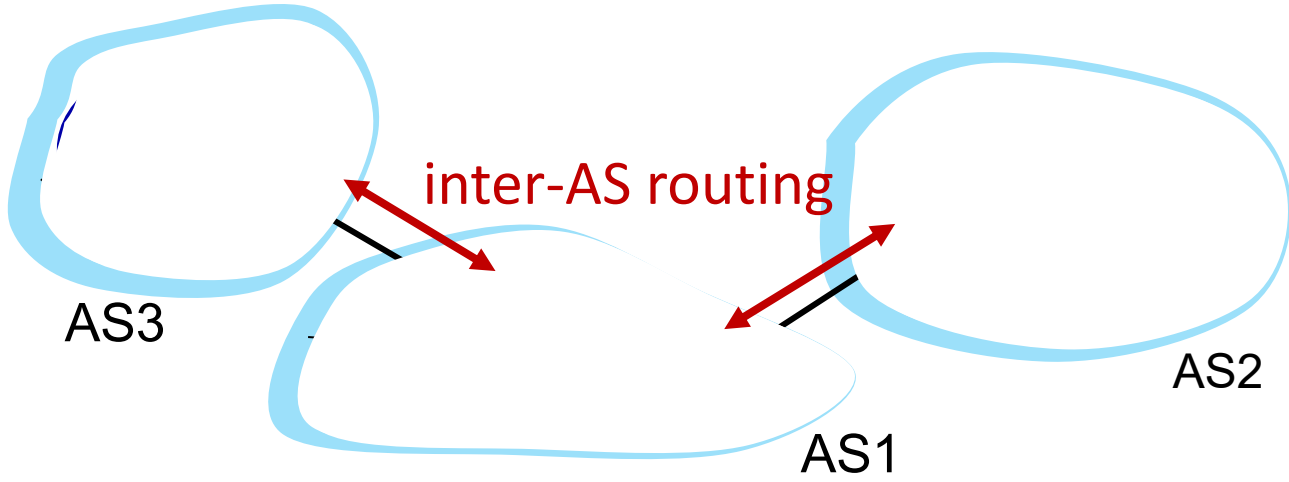
Router di dorsale (*backbone router*): esegue OSPF limitatamente alla dorsale

Livello di rete: tabella di marcia del “piano di controllo”

- introduzione
- algoritmi di instradamento
  - link state
  - distance vector
- instradamento interno al sistema autonomo: OSPF
- instradamento tra sistemi autonomi: BGP
- piano di controllo SDN
- Internet Control Message Protocol

- gestione e configurazione della rete
  - SNMP
  - NETCONF/YANG

Sistemi autonomi (AS) interconnessi



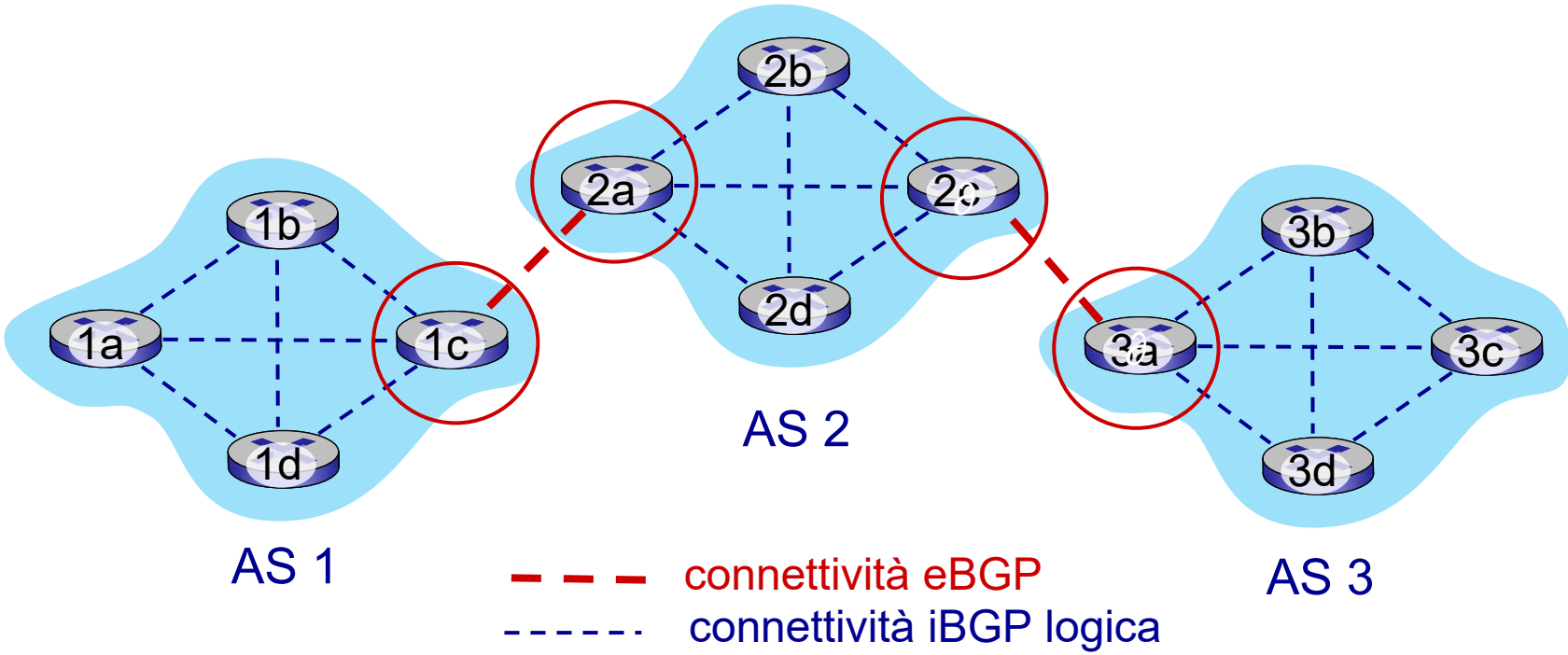
intra-AS (o “intra-domain”): instradamento tra router *dentro lo stesso AS (“rete”)*

inter-AS (o “inter-domain”): instradamento *tra* AS

Instradamento Internet inter-AS: BGP

- BGP (Border Gateway Protocol): *il* protocollo di fatto per l'instradamento inter-domain
  - "colla che tiene insieme Internet"
- permette alla sottorete di pubblicizzare la sua esistenza e le destinazioni che può raggiungere al resto di Internet: *"Io sono qui, ecco chi posso raggiungere e come"*
- BGP fornisce a ciascun AS un mezzo per:
  - ottenere informazioni sulla raggiungibilità dei prefissi di sottorete da parte dei sistemi confinanti (eBGP)
  - determinare le rotte verso altre reti sulla base delle informazioni di raggiungibilità e di *politiche (policy)*
  - propagare le informazioni di raggiungibilità a tutti i router interni all'AS (iBGP)
  - annunciare (alle reti confinanti) le informazioni sulla raggiungibilità delle destinazioni

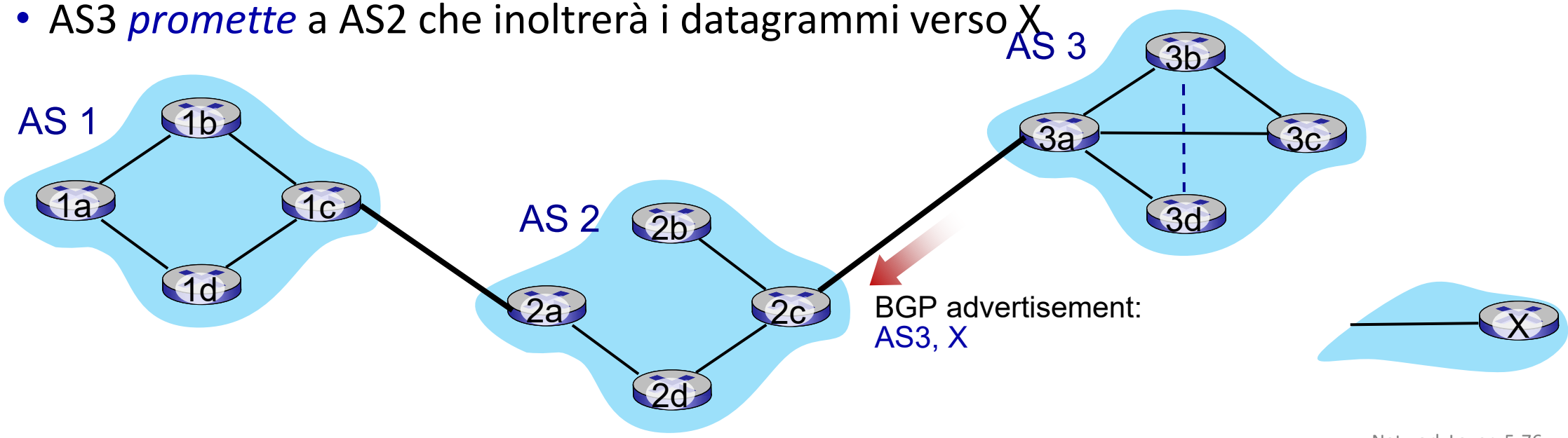
Connessioni eBGP, iBGP



router gateway eseguono sia il protocollo eBGP sia il protocollo iBGP

Nozioni di base su BGP

- Sessione BGP: due router BGP (“peers”) si scambiano messaggi BGP attraverso un connessione TCP semi-permanente:
  - annunciare *percorsi* verso diversi prefissi di rete di destinazione (BGP è un protocollo “path vector”)
- Quando il gateway 3a di AS3 annuncia il percorso AS3,X al gateway 2c di AS2:
  - AS3 *promette* a AS2 che inoltrerà i datagrammi verso X



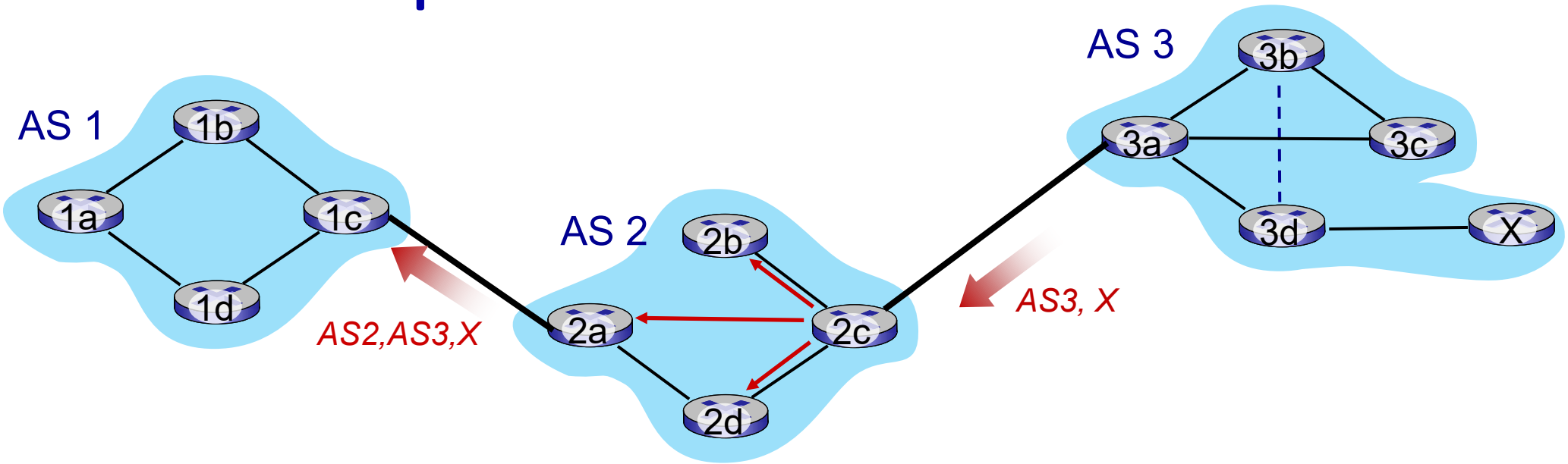
Messaggi del protocollo BGP

- I messaggi BGP sono scambiati tra peer su connessioni TCP
- Messaggi BGP [RFC 4371]:
  - OPEN: apre la connessione TCP al peer BGP remoto e autentica il peer BGP mittente
  - **UPDATE:** annuncia un nuovo percorso (o ritira il vecchio)
  - KEEPALIVE: mantiene in vita la connessione in assenza di UPDATE; inoltre ACK della richiesta OPEN
  - NOTIFICATION: segnala gli errori nel messaggio precedente; viene usato anche per chiudere la connessione

Attributi dei percorsi e rotte BGP

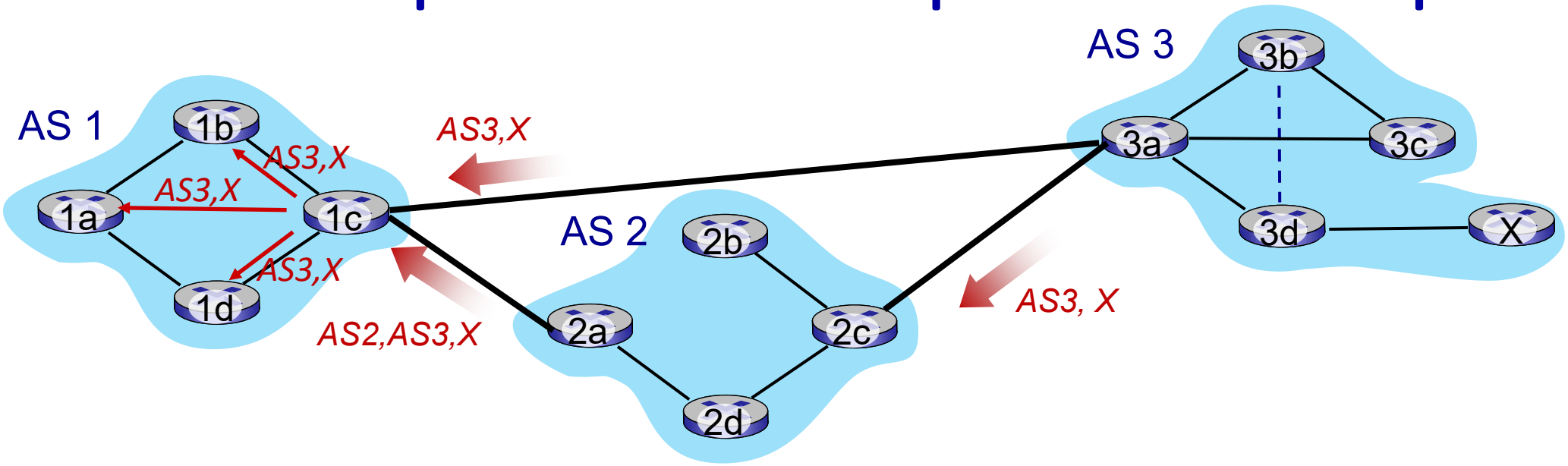
- Rotta (*route*) annunciata da BGP: prefisso + attributi
  - prefisso: la destinazione che viene annunciata
  - due attributi importanti:
    - AS-PATH: elenco degli AS attraverso i quali è passato l'annuncio del prefisso
    - NEXT-HOP: indirizzo IP dell'interfaccia del router che inizia l'AS-PATH
- instradamento basato su politiche:
  - Un gateway che riceve un annuncio di percorso usa una *import policy* per accettare/declinare il percorso (es., mai instradare attraverso AS Y).
  - Le politiche dell'AS determinano anche se *annunciare* un percorso a altri AS vicini

Annuncio di percorso BGP



- il router 2c in AS2 riceve l'annuncio del percorso AS3,X (attraverso eBGP) dal router 3a in AS3
- sulla base delle politiche di AS2, il router 2c in AS2 accetta il percorso AS3,X, e lo propaga (attraverso iBGP) a tutti i router in AS2
- sulla base delle politiche di AS2, il router 2a in AS2 annuncia (attraverso eBGP) il percorso AS2, AS3, X al router 1c in AS1

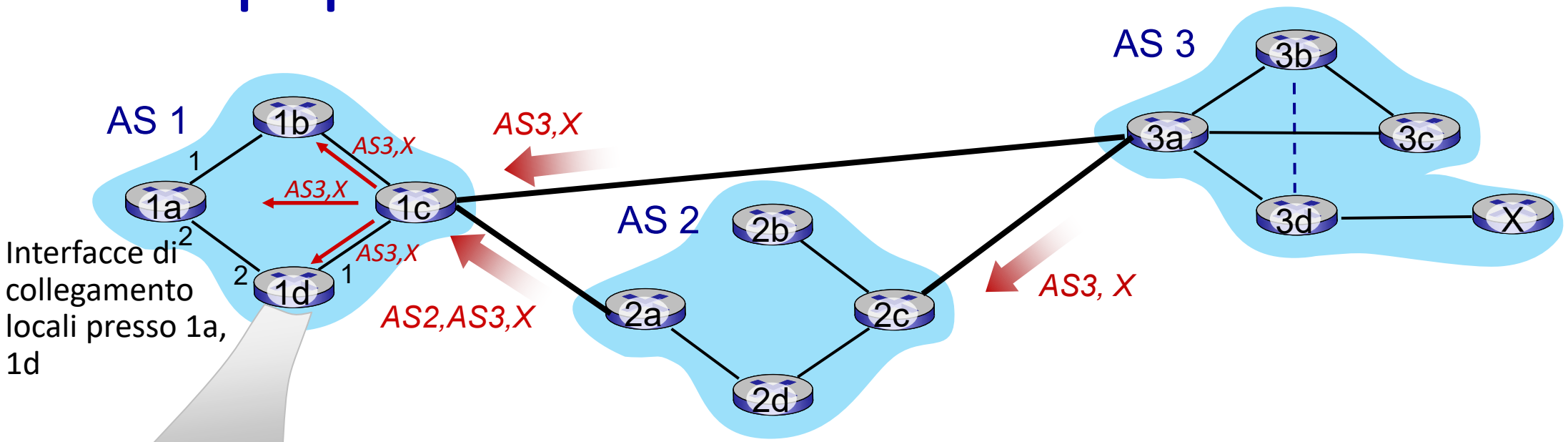
Annuncio di percorso BGP: percorsi multipli

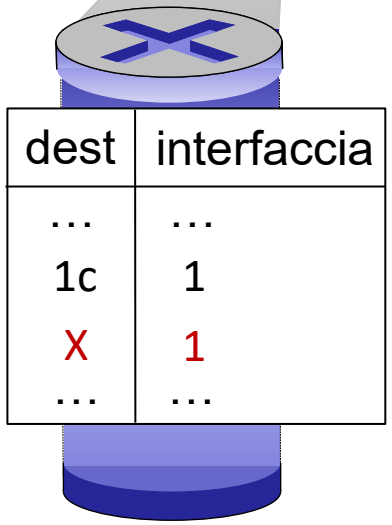


un router gateway potrebbe venire a conoscenza di percorsi molteplici verso una certa destinazione:

- il router gateway 1c di AS1 apprende il percorso *AS2,AS3,X* da 2a
- il router gateway 1c di AS1 apprende il percorso *AS3,X* a 3a
- sulla base di *politiche,* il router gateway 1c in AS1 sceglie il percorso *AS3,X* e annuncia il percorso dentro l'AS attraverso iBGP

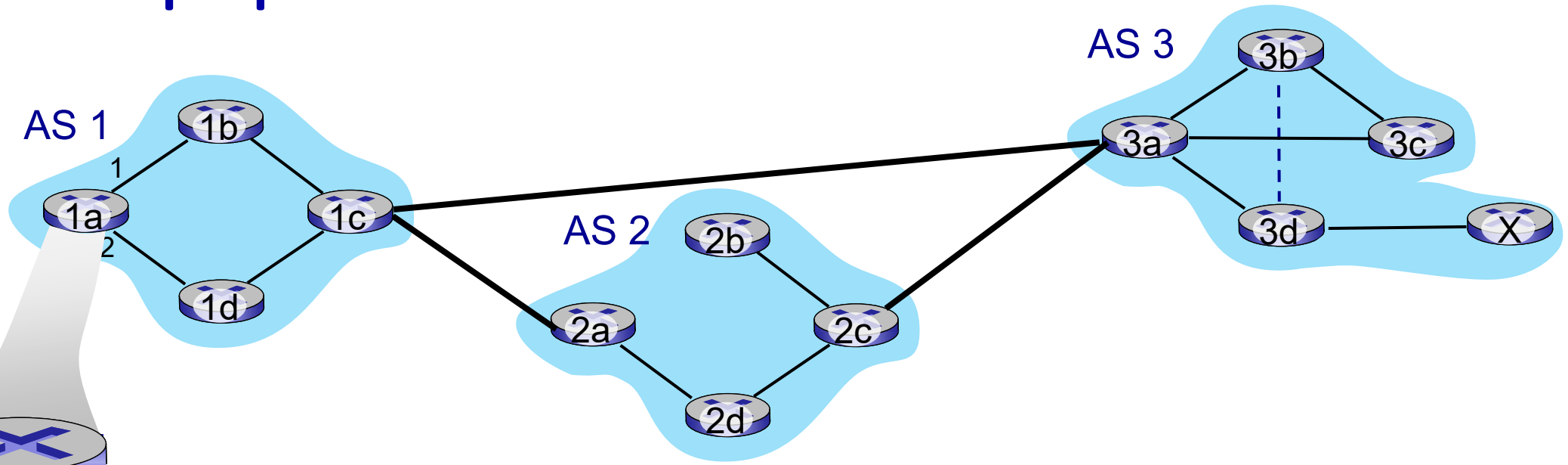
BGP: popolare le tabelle di inoltro

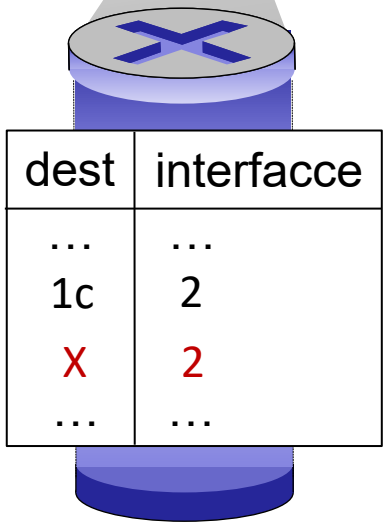




- ricordare: 1a, 1b, 1d apprendono tranite iBGP da 1c: “il percorso verso X passa attraverso 1c”
- presso 1d: OSPF intra-domain routing: per raggiungere 1c, usa l'interfaccia 1
- presso 1d: per raggiungere X, usa l'interfaccia 1

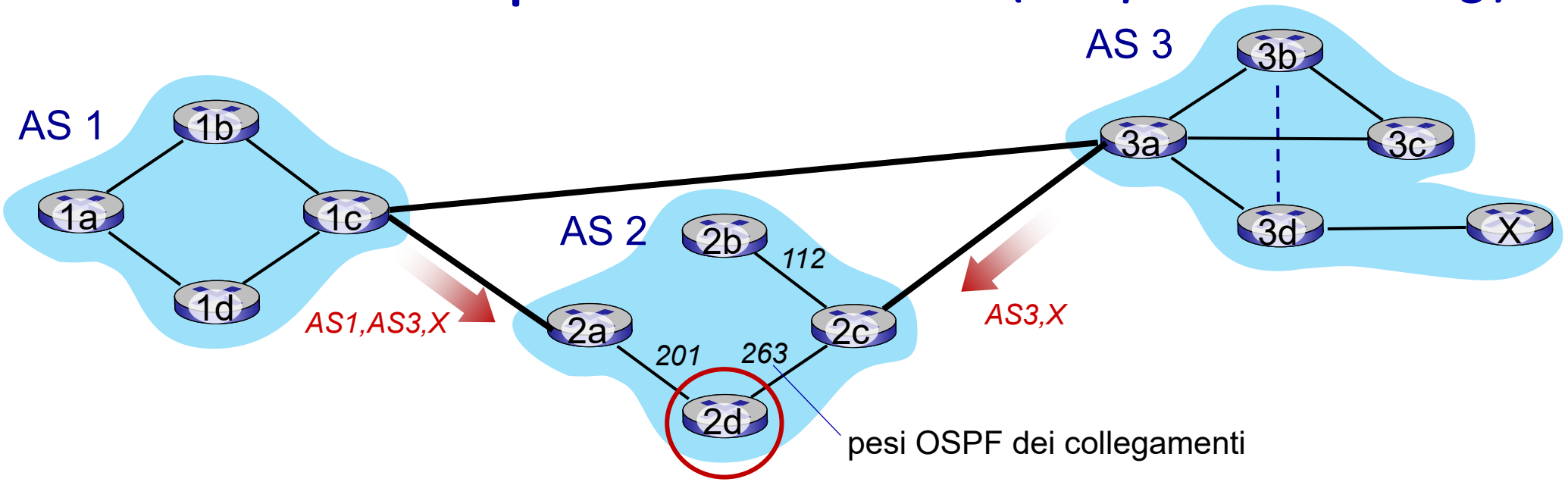
BGP: popolare le tabelle di inoltro





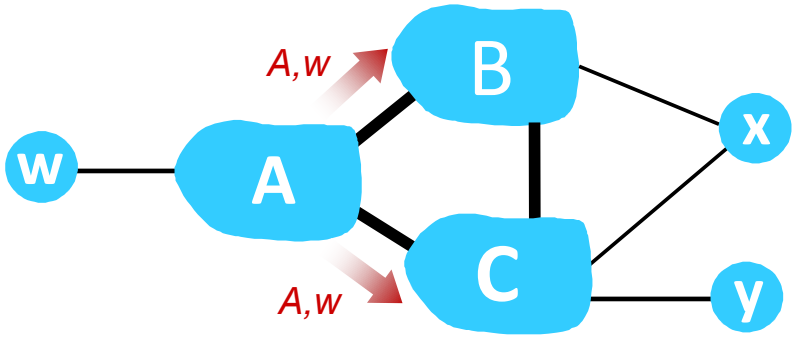
- ricordare: 1a, 1b, 1d apprendono tranite iBGP da 1c: “il percorso verso X passa attraverso 1c”
- presso 1d: OSPF intra-domain routing: per raggiungere 1c, usa l'interfaccia 1
- presso 1d: per raggiungere X, usa l'interfaccia 1
- presso 1a: OSPF intra-domain routing: per raggiungere 1c, usa l'interfaccia 2
- presso 1a: per raggiungere X, usa l'interfaccia 2

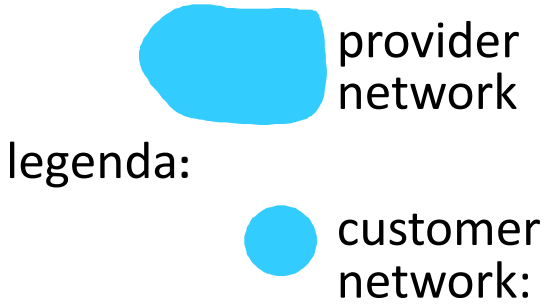
Instradamento a patata bollente (*hot potato routing*)



- 2d apprende (tramite iBGP) che può instradare verso X via 2a o 2c
- instradamento a patata bollente: sceglie il gateway locale che ha il minimo costo *intra-AS* (es., 2d sceglie 2a, nonostante il maggior numero di hop hops verso *X*): non preoccupatevi del costo inter-AS!

BGP: implementare le politiche attraverso gli annunci

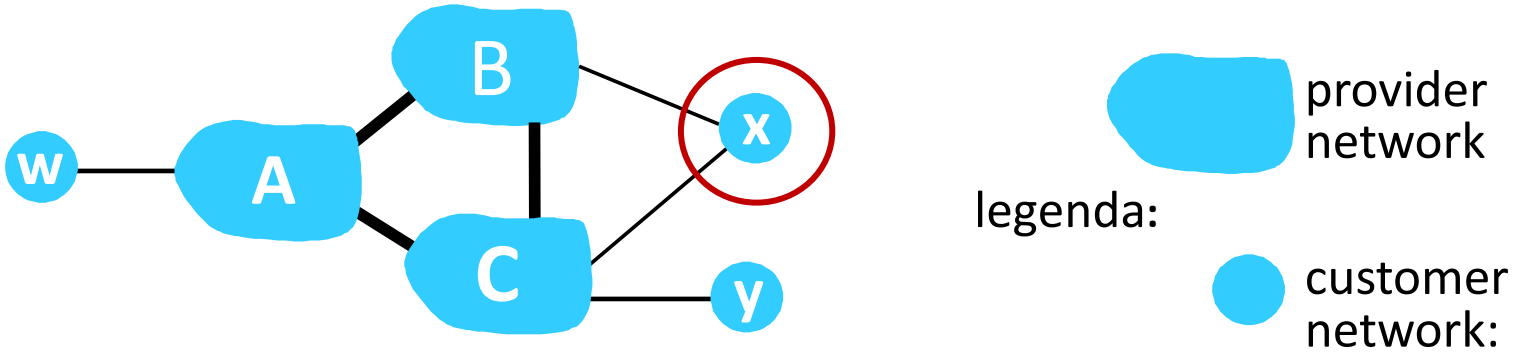




L'ISP vuole instradare il traffico solo verso/da le reti dei propri clienti (non vuole trasportare il traffico di transito tra altri ISP - una politica tipica del "mondo reale")

- A annuncia il percorso Aw a B e a C
- B *scegli di non annunciare* BAw a C!
  - B non riceve alcuna "entrata" per l'instradamento CBAw, visto che né C, A, w sono clienti di B
  - C *non* viene a conoscenza del percorso CBAw
- C instraderà CAw (non usando B) per raggiungere w

BGP: implementare le politiche attraverso gli annunci (+)



L'ISP vuole instradare il traffico solo verso/da le reti dei propri clienti (non vuole trasportare il traffico di transito tra altri ISP - una politica tipica del “mondo reale”)

- A,B,C sono provider network
- x,w,y sono customer (delle provider networks)
- x è dual-homed: connessa a due reti
- *politica da applicare:* x non vuole instradare da B a C attraverso x
  - .. quindi x non annuncerà a B un percorso verso C

Selezione delle rotte BGP

- Il router può conoscere più di un percorso verso l'AS di destinazione, seleziona il percorso in base a:
  1. valore dell'attributo di **preferenza locale**: decisione politica
  2. AS-PATH più breve
  3. router NEXT-HOP più vicino: instradamento a patata bollente
  4. identificatori BGP

Perché diversi instradamenti Intra- e Inter-AS?

politiche:

- inter-AS: l'amministratore vuole avere il controllo sul modo in cui viene instradato il suo traffico, su chi passa attraverso la sua rete
- intra-AS: singolo amministratore, quindi le politicche sono meno problematiche

scalabilità:

- il routing gerarchico consente di ridurre le dimensioni delle tabelle e il traffico di aggiornamento.

prestazioni:

- intra-AS: può concentrarsi sulle prestazioni
- inter-AS: le politiche sono dominanti rispetto alla prestazioni

Livello di rete: tabella di marcia del “piano di controllo”

- introduzione
- algoritmi di instradamento
  - link state
  - distance vector
- instradamento interno al sistema autonomo: OSPF
- instradamento tra sistemi autonomi: BGP
- piano di controllo SDN
- Internet Control Message Protocol

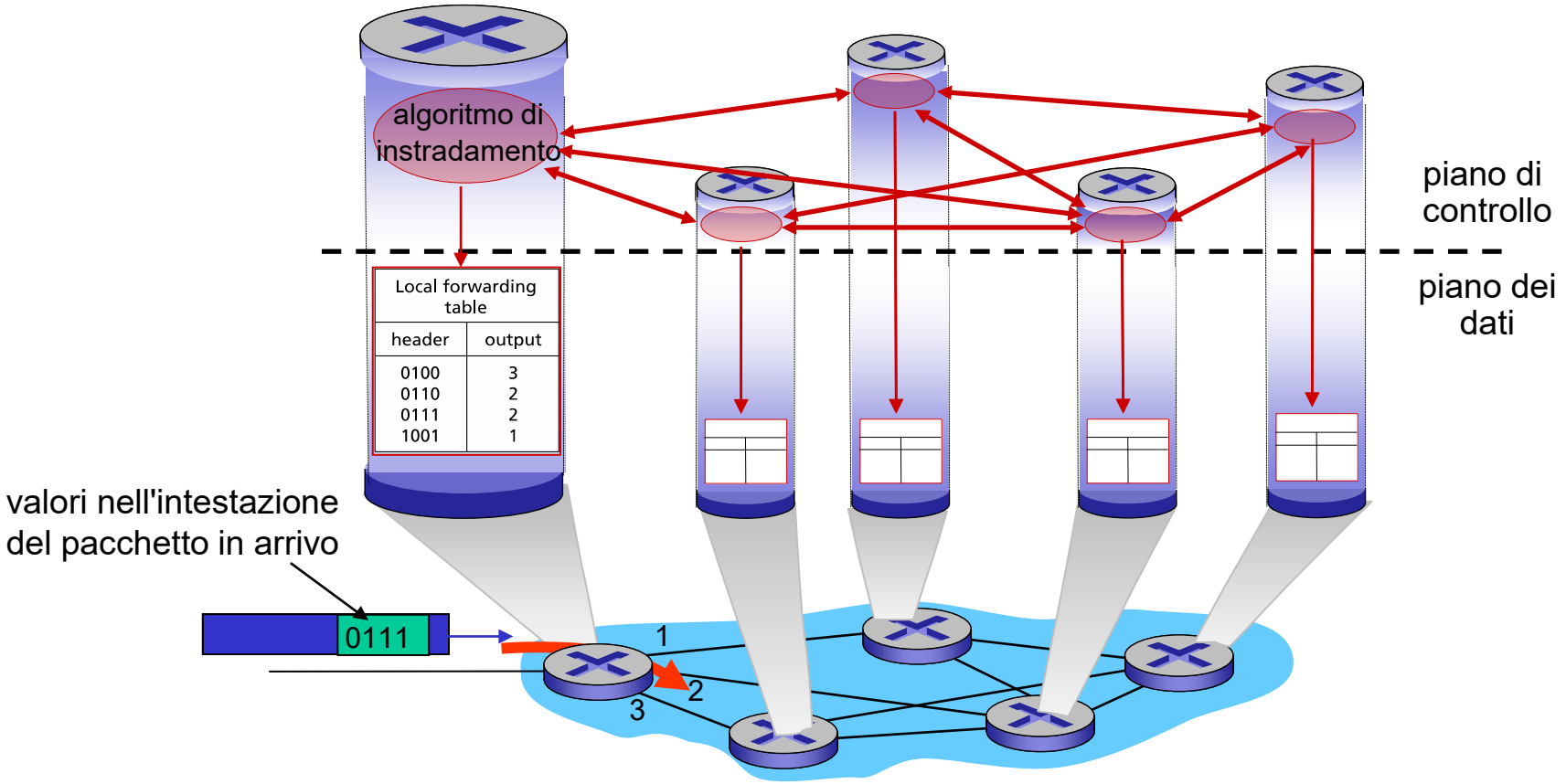
- gestione e configurazione della rete
  - SNMP
  - NETCONF/YANG

Software defined networking (SDN)

- Livello di rete di Internet: storicamente implementato tramite un approccio di controllo distribuito e per router:
  - Un *router monolitico* contiene l'hardware di commutazione (*switching*), esegue una implementazione proprietaria dei protocolli standard di Internet (IP, RIP, IS-IS, OSPF, BGP) in un sistema operativo proprietario specializzato per dispositivi di rete (es. Cisco IOS)
  - "middlebox" differenti per differenti funzioni del livello di rete: firewalls, load balancers, NAT, ..
- ~2005: rinnovato interesse nel ripensare il piano di controllo della rete

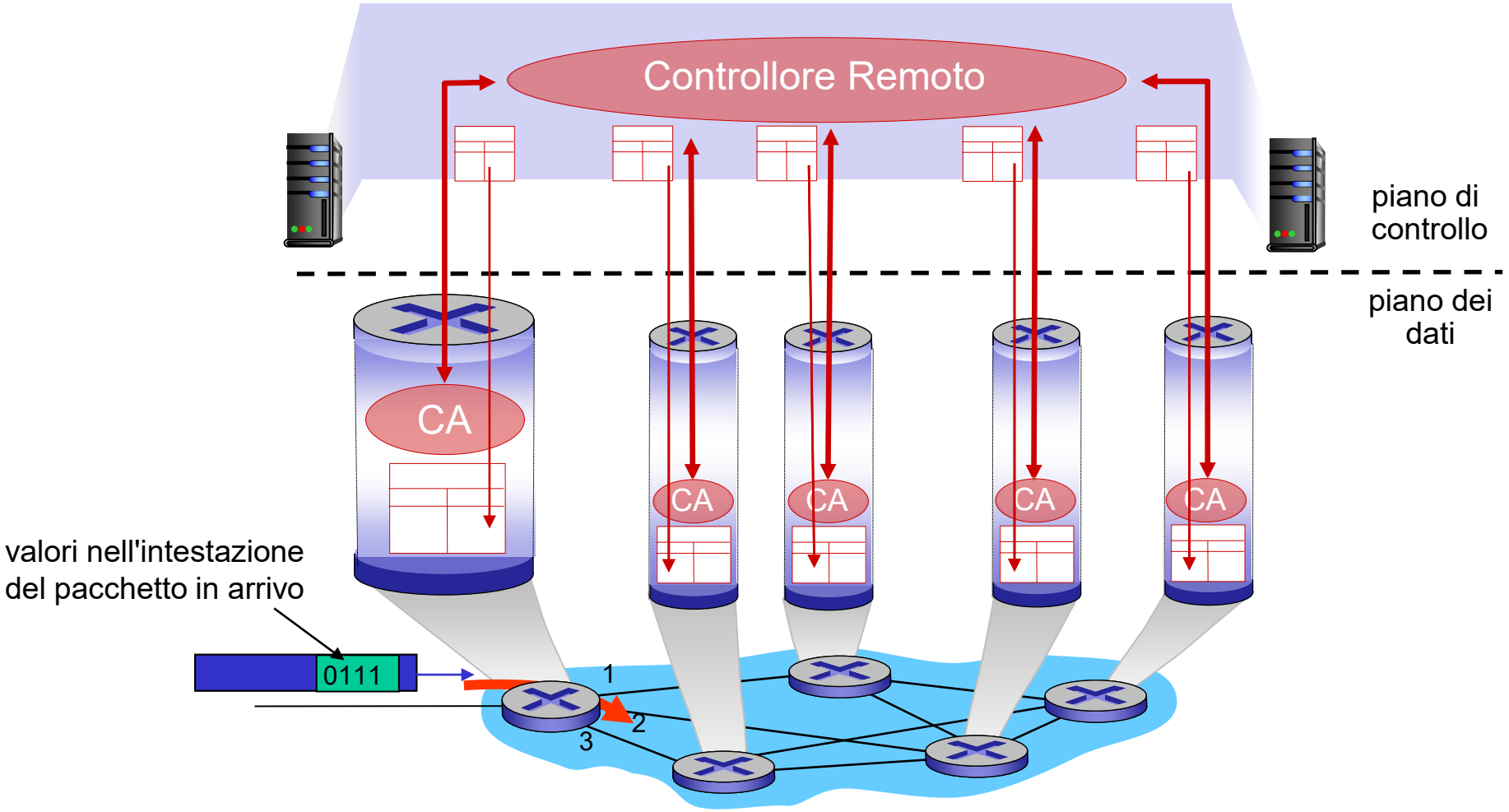
Piano di controllo per rotuer

I singoli componenti dell'algoritmo di instradamento *in ogni router* interagiscono nel piano di controllo.



Piano di controllo Software-Defined Networking (SDN)

Il controller remoto calcola e installa le tabelle di inoltro nei router.



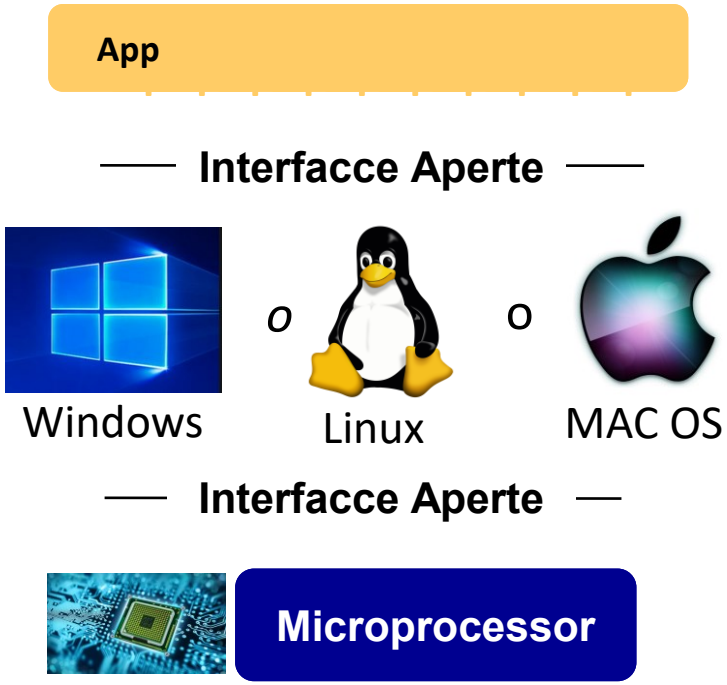
Software defined networking (SDN)

Perché un piano di controllo *logicamente centralizzato*?

- gestione più semplice della rete: evitare errori di configurazione dei router, maggiore flessibilità dei flussi di traffico
- inoltro basato su tabelle (ricordate OpenFlow API) permette la "programmazione" dei router
  - la “programmazione” centralizzata è più semplice: calcola le tabelle centralmente e poi distribuiscile
  - la “programmazione” distribuita è più difficile: calcolo delle tabelle come risultato di un algoritmo (protocollo) distribuito implementato in ogni singolo router
- implementazione aperta (non proprietaria) del piano di controllo
  - promuovere l'innovazione

Analogia con l'SDN: dal mainframe alla rivoluzione del PC

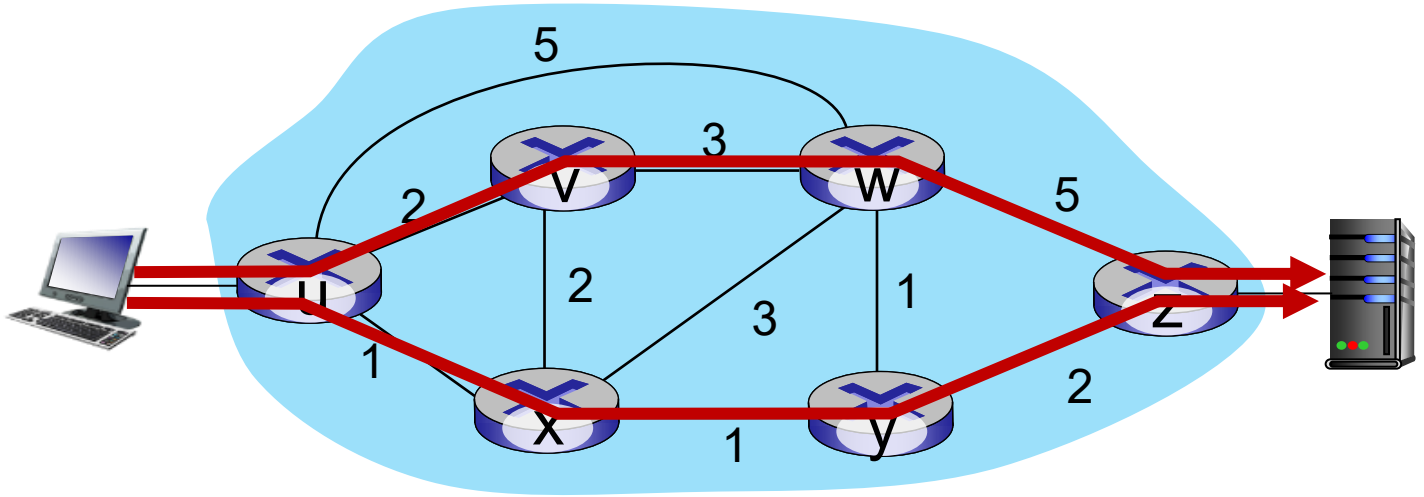
Integrato verticalmente
chiuso, proprietario
Innovazione lenta
Una piccola industria



Orizzontale
Interfacce aperte
Innovazione rapida
Un'industria enorme

* Slide courtesy: N. McKeown

Ingegneria del traffico: difficile con il routing tradizionale

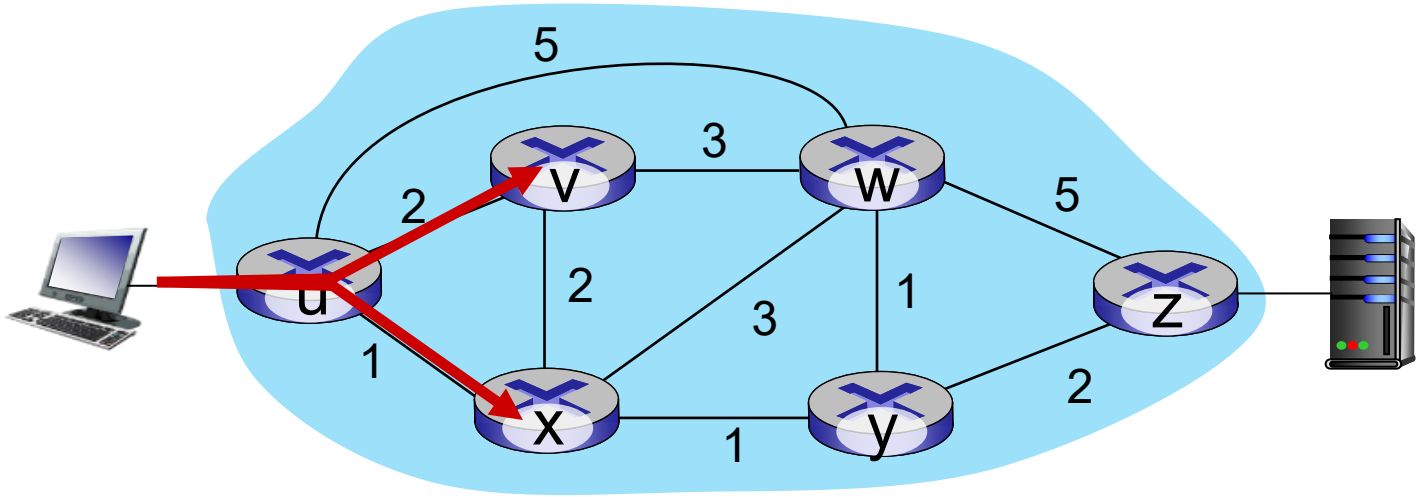


D: cosa succede se l'operatore di rete vuole che il traffico da *u* a *z* fluisca lungo *uvwz*, anziché *uxyz*?

R: è necessario ridefinire i pesi dei collegamenti in modo che l'algoritmo di instradamento del traffico calcoli le rotte di conseguenza (o necessitiamo di un nuovo algoritmo di instradamento)!

I pesi dei collegamenti sono le solo "manopole" di controllo: non c'è molto controllo!

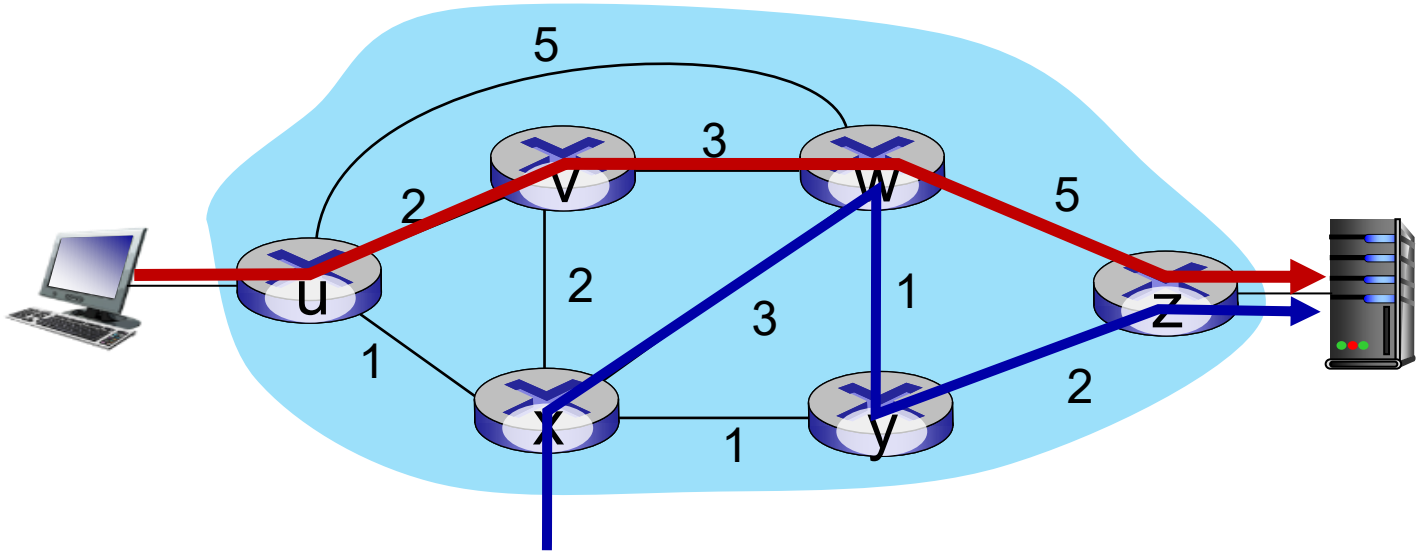
Ingegneria del traffico: difficile con il routing tradizionale



*D: c*osa succede se l'operatore di rete vuole dividere il traffico *u*-to-*z* lungo *uvwz e uxyz* (bilanciamento del carico)?

R: non può farlo (o ha bisogno di un nuovo algoritmo di routing)

Ingegneria del traffico: difficile con il routing tradizionale

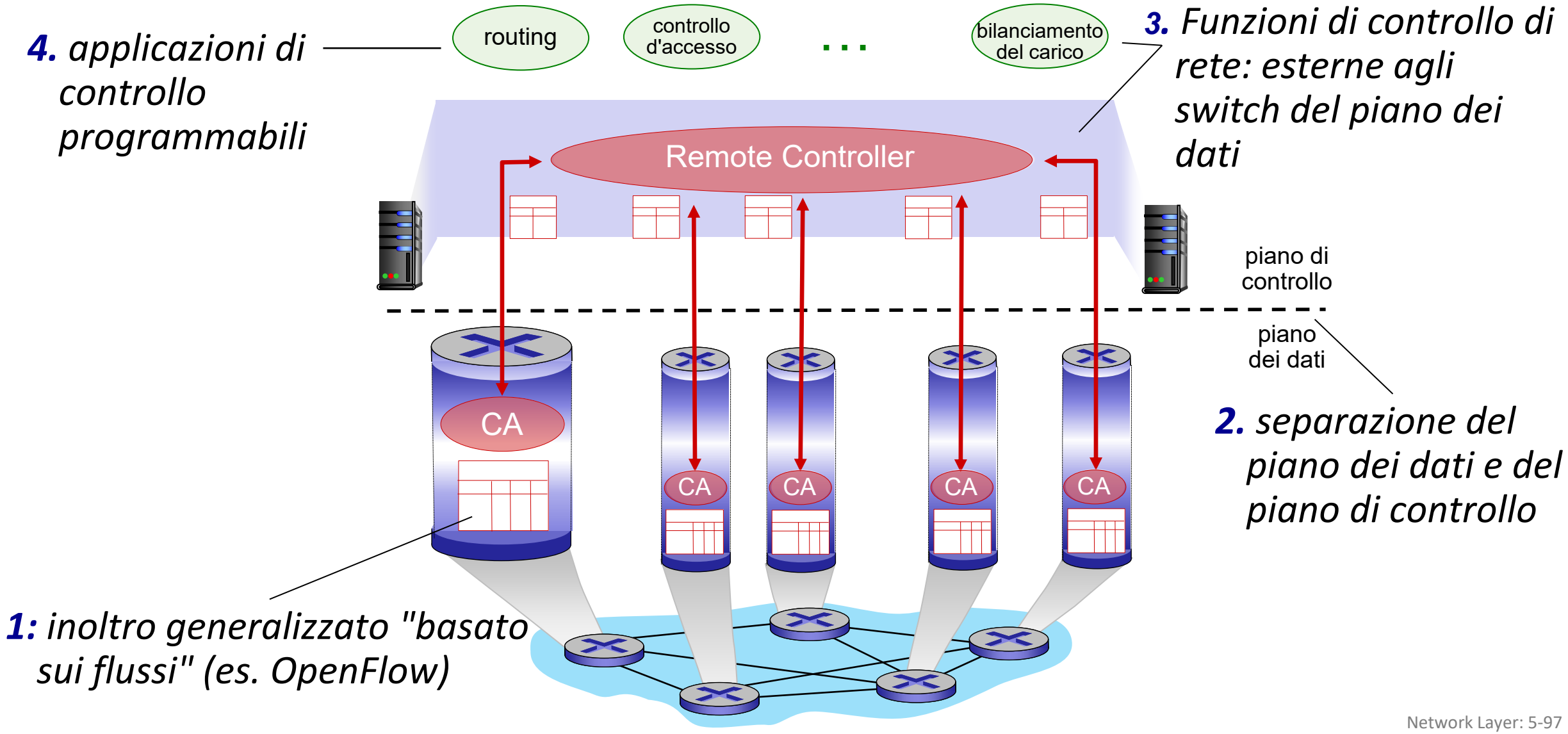


D: e se *w* volesse instradare il traffico blu e rosso in modo diverso da *w* a *z*?

R: non può farlo (con l'inoltro basato sulla destinazione e l'instradamento LS e DV).

Abbiamo appreso che l'inoltro generalizzato e l'SDN possono essere usati per raggiungere *qualsiasi* instradamento si desideri

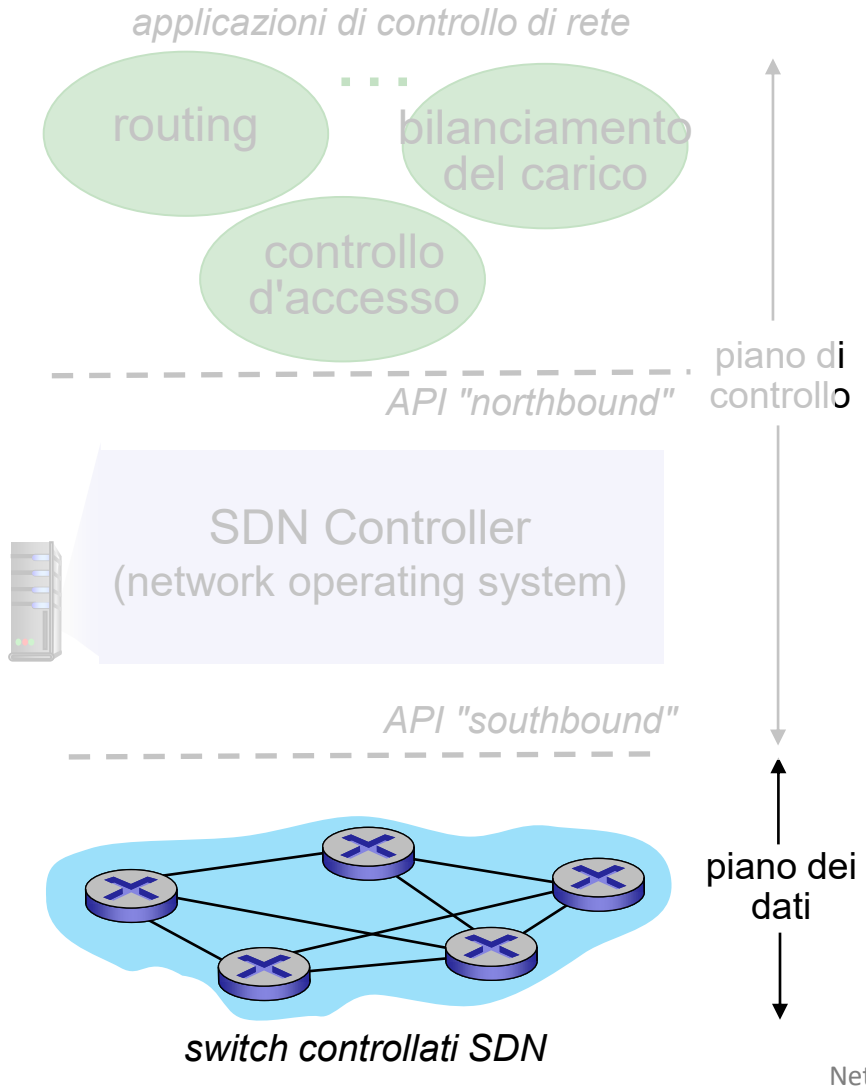
Software defined networking (SDN)



Software defined networking (SDN)

Switch del piano dei dati :

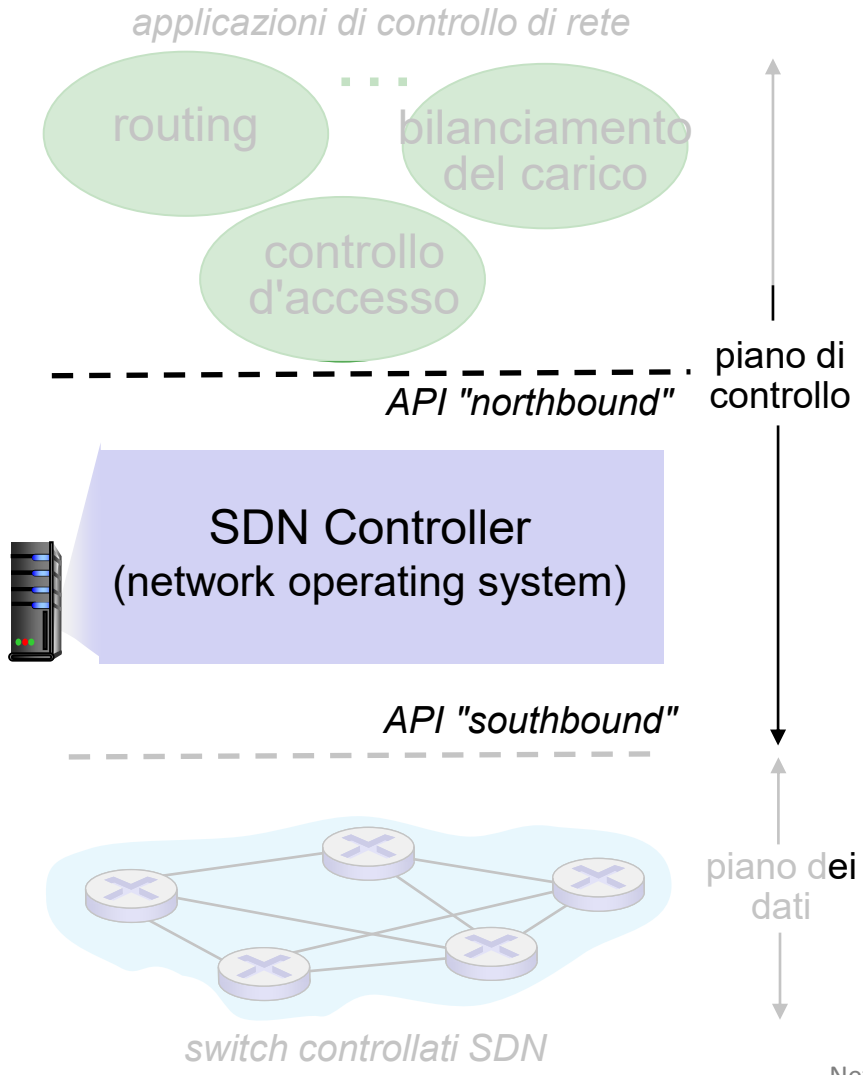
- switch veloci e semplici che implementano l'inoltro generalizzato del piano dei dati in hardware
- tabella dei flussi (inoltro) calcolata, installata sotto la supervisione del controllore
- API per il controllo degli switch basato su tabelle (es., OpenFlow)
  - definisce ciò che è controllabile e ciò che non lo è
- protocollo di comunicazione con il controllore (es. OpenFlow)



Software defined networking (SDN)

SDN controller (network OS):

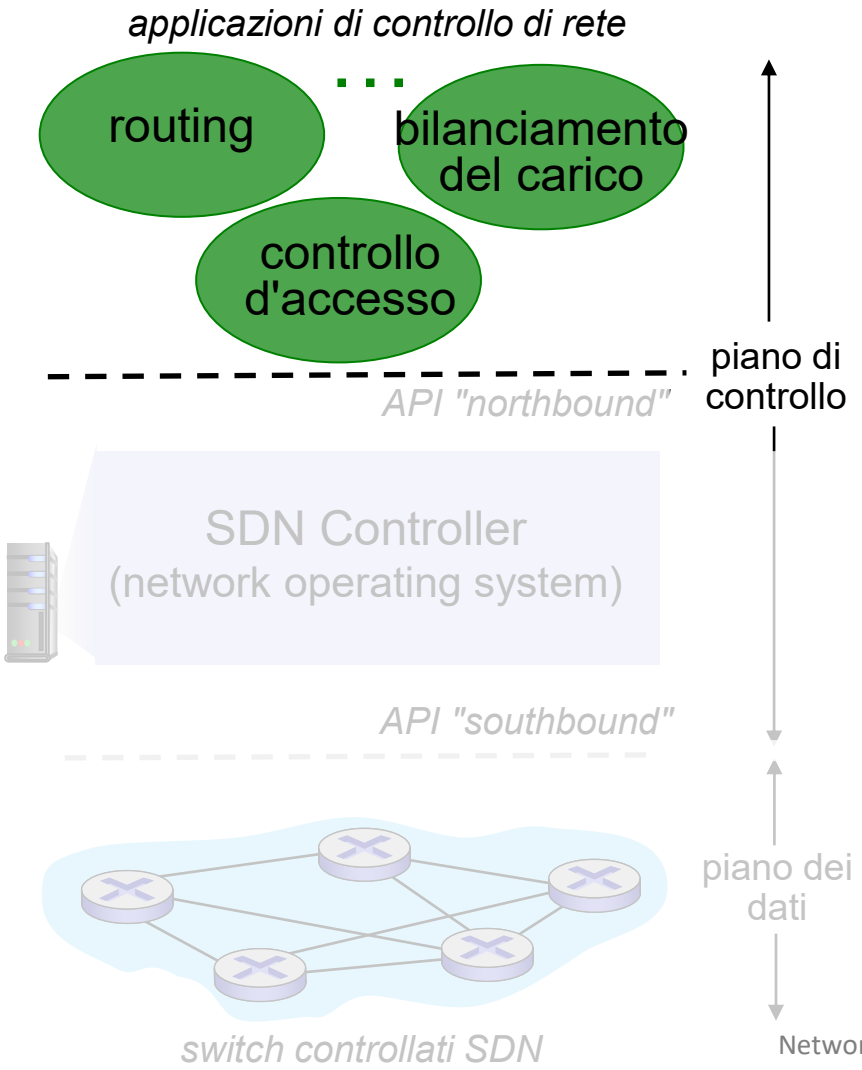
- mantiene le informazioni sullo stato della rete
- interagisce con le applicazioni di controllo della rete “in alto” tramite API "northbound"
- interagisce con gli switch di rete “in basso” tramite API "southbound"
- implementato come sistema distribuito per garantire prestazioni, scalabilità, tolleranza ai guasti, robustezza e sicurezza.



Software defined networking (SDN)

Applicazioni di controllo di rete:

- “cervelli” di controllo: implementano le funzioni di controllo utilizzando servizi di livello inferiore attraverso API fornite dal controller SDN
- *scorporate:* può essere fornito da terzi: distinto dal fornitore di routing o dal controller SDN

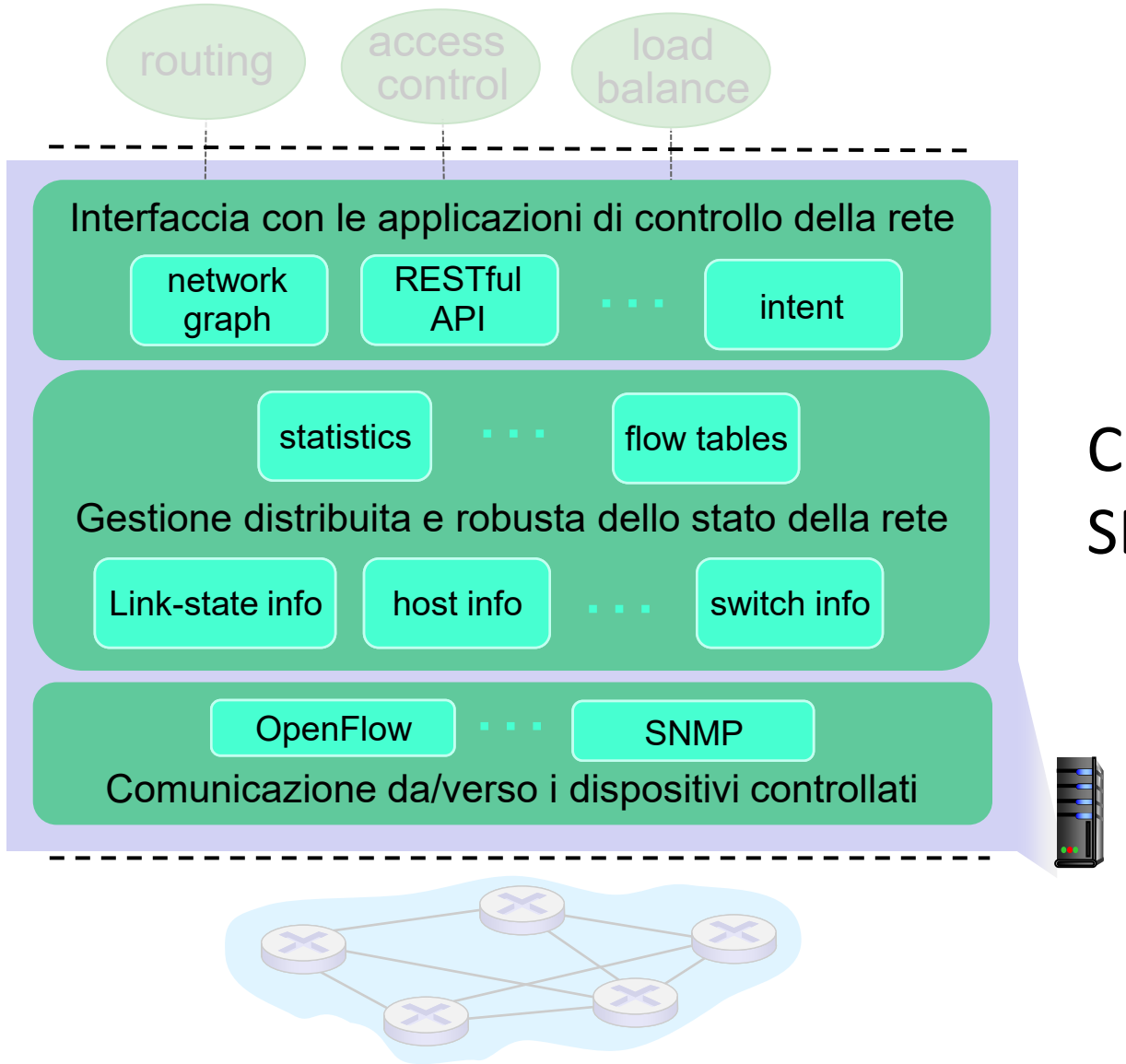


Componenti di un Controller SDN

livello di interfaccia con le applicazioni di controllo della rete: astrazioni/API

gestione dello stato della rete: stato dei collegamenti di rete, degli switch, dei servizi: un *database distribuito*

comunicazione: comunicazione tra il controller SDN e gli switch controllati

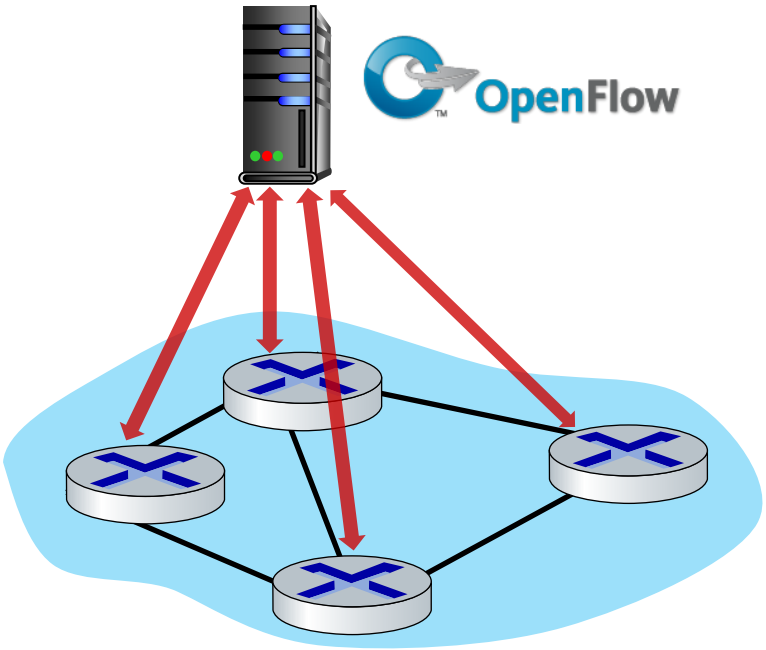


Controller SDN

Protocollo OpenFlow

- opera tra controllore e switch
- TCP utilizzato per lo scambio di messaggi
  - crittografia opzionale
- tre classi di messaggi OpenFlow:
  - controller-to-switch
  - asynchronous (switch to controller)
  - symmetric (misc.)
- distinta dall'API OpenFlow
  - API utilizzata per specificare azioni di inoltro generalizzate

Controller OpenFlow

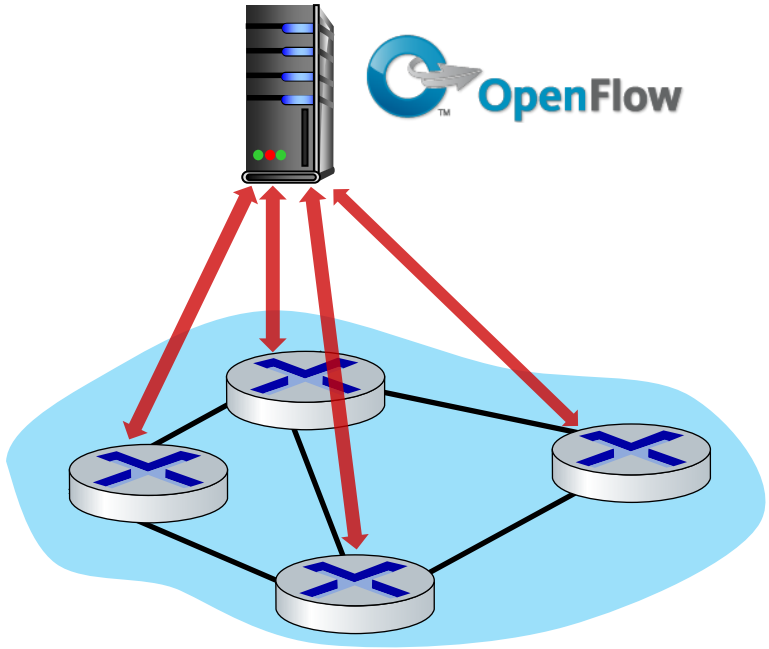


OpenFlow: messaggi controller-to-switch

Messaggi chiave controller-to-switch

- *features:* il controllore interroga le caratteristiche dello switch, lo switch risponde
- *configure:* il controllore interroga/imposta i parametri di configurazione dello switch
- *modify-state:* aggiungere, eliminare, modificare voci di flusso nelle tabelle OpenFlow
- *packet-out:* Il controllore può inviare questo pacchetto da una specifica porta dello switch

Controller OpenFlow

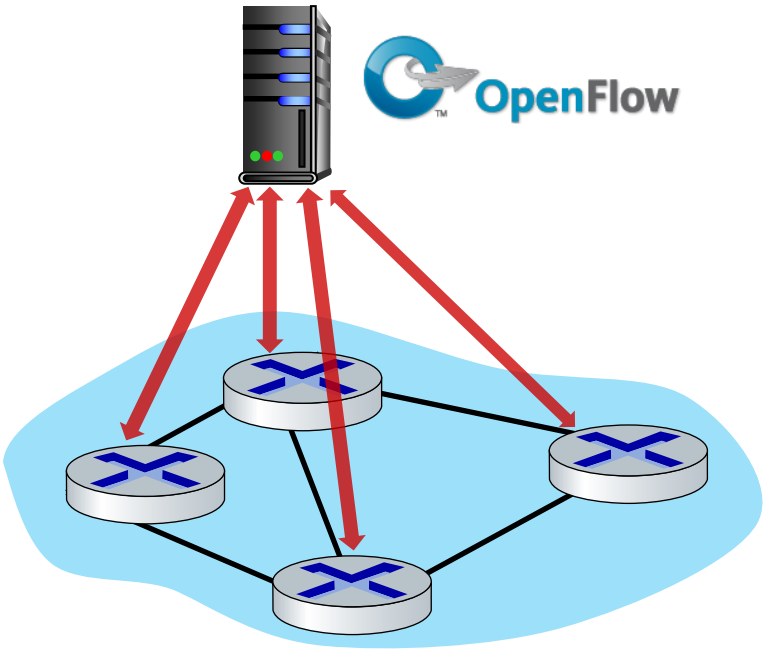


OpenFlow: messaggi switch-to-controller

Messaggi chiave switch-to-controller

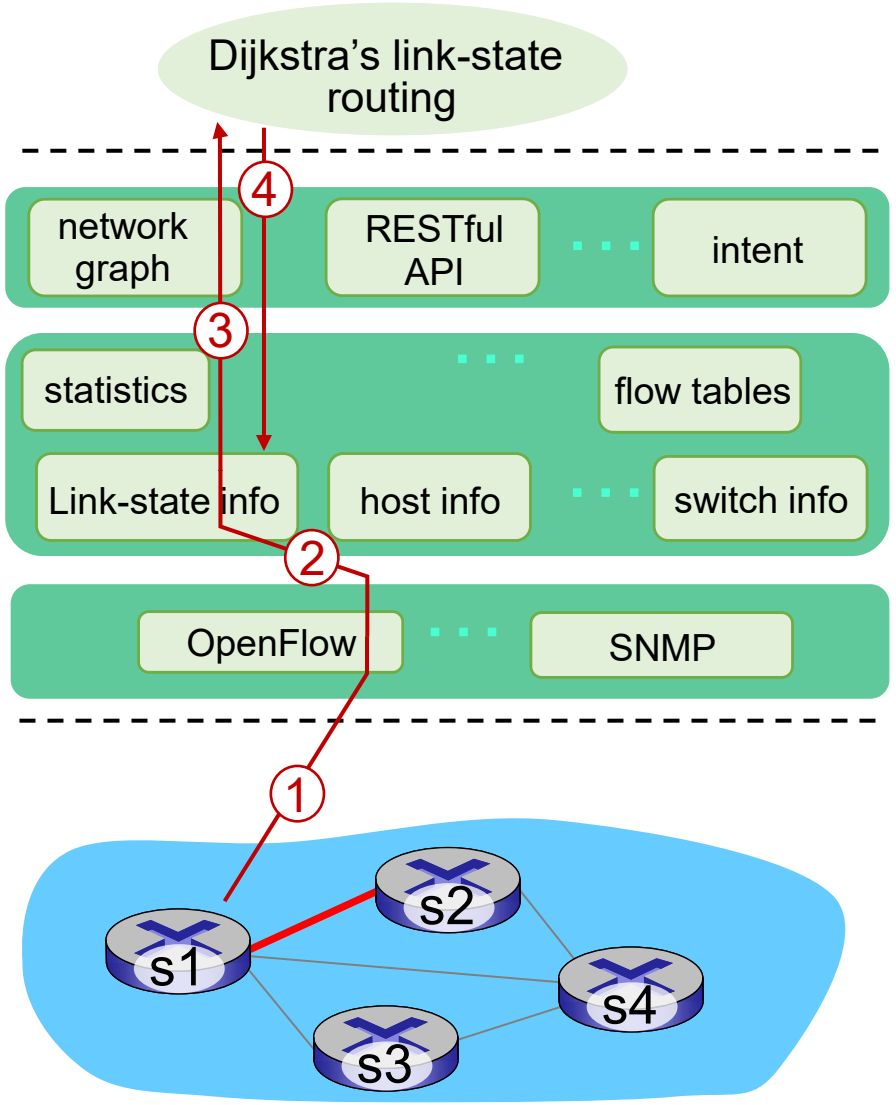
- *packet-in:* trasferire il pacchetto (e il relativo controllo) al controllore. Vedere il messaggio packet-out dal controllore
- *flow-removed:* voce della tabella di flusso cancellata nello switch
- *port status:* informare il controllore di una modifica su una porta.

Controller OpenFlow



Fortunatamente, gli operatori di rete non “programmano” gli switch creando/inviando direttamente messaggi OpenFlow. Utilizzano invece un'astrazione di livello superiore a livello di controller

SDN: Esempio di interazione tra piano dei dati e piano di controllo



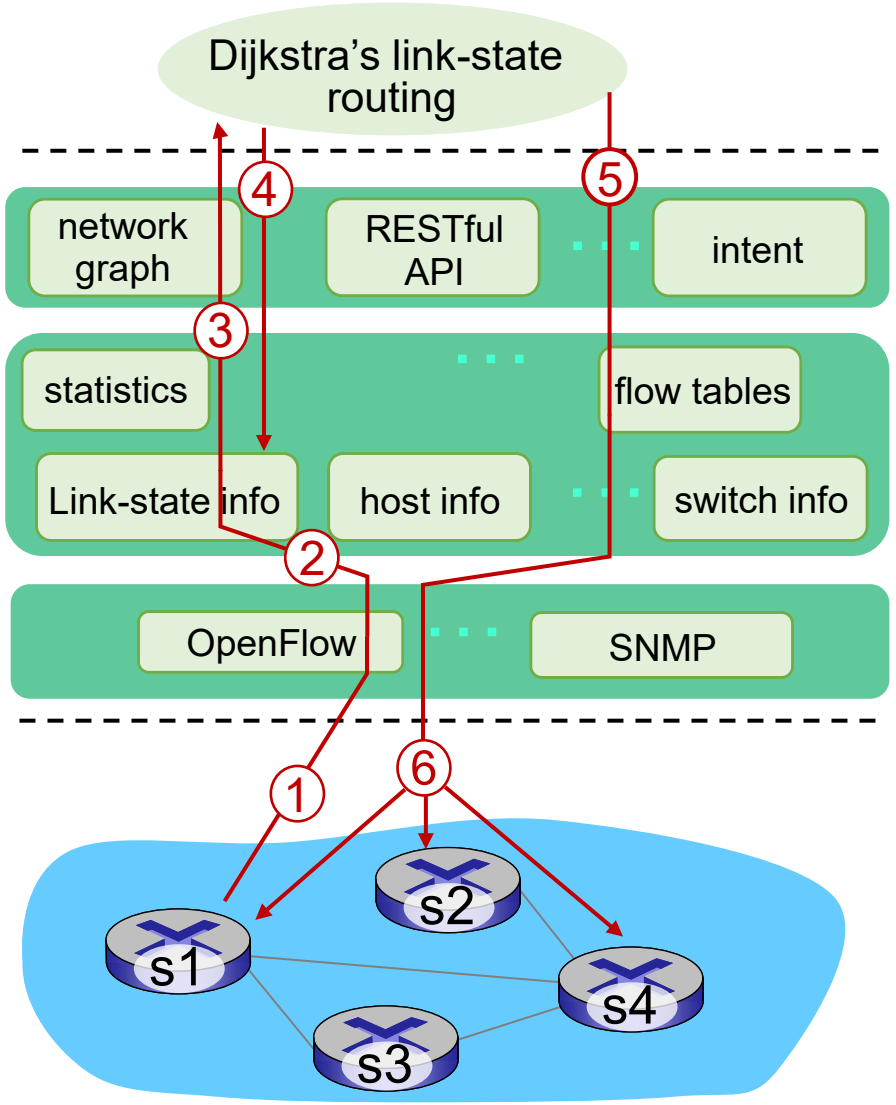
① S1, a causa di un guasto del collegamento, utilizza il messaggio di stato della porta OpenFlow per notificare il controllore.

② Il controller SDN riceve il messaggio OpenFlow, aggiorna le informazioni sullo stato del collegamento

③ L'applicazione dell'algoritmo di routing di Dijkstra si è registrata in precedenza per essere richiamata quando lo stato dei collegamenti cambia. Viene chiamata.

④ L'algoritmo di routing di Dijkstra accede alle informazioni sul grafo della rete, alle informazioni sullo stato dei collegamenti nel controllore e calcola nuovi percorsi.

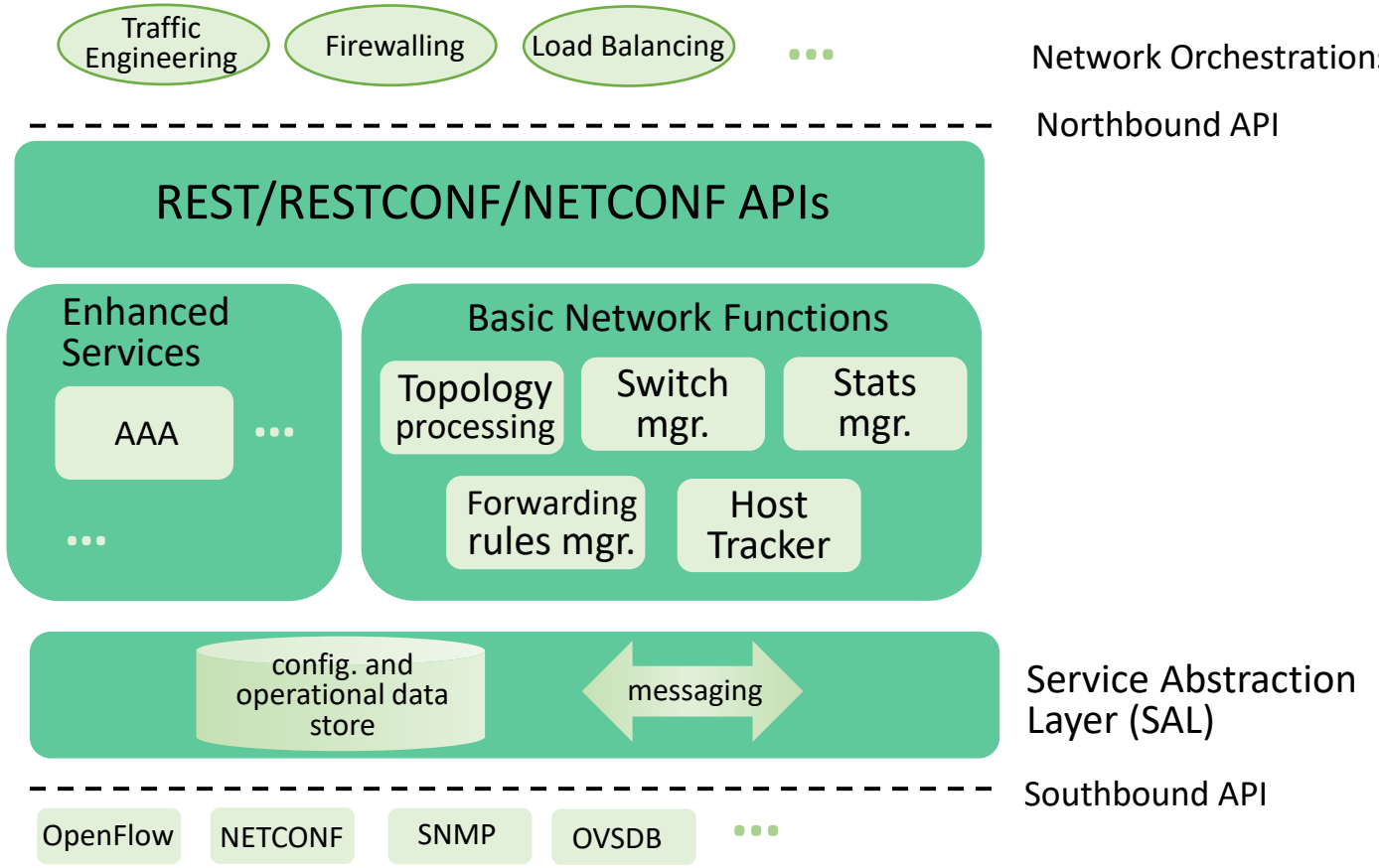
SDN: Esempio di interazione tra piano dei dati e piano di controllo



⑤ l'applicazione di link state routing interagisce con il componente flow-table-computation del controller SDN, che calcola le nuove tabelle di flusso necessarie.

⑥ il controllore utilizza OpenFlow per installare nuove tabelle negli switch che necessitano di un aggiornamento

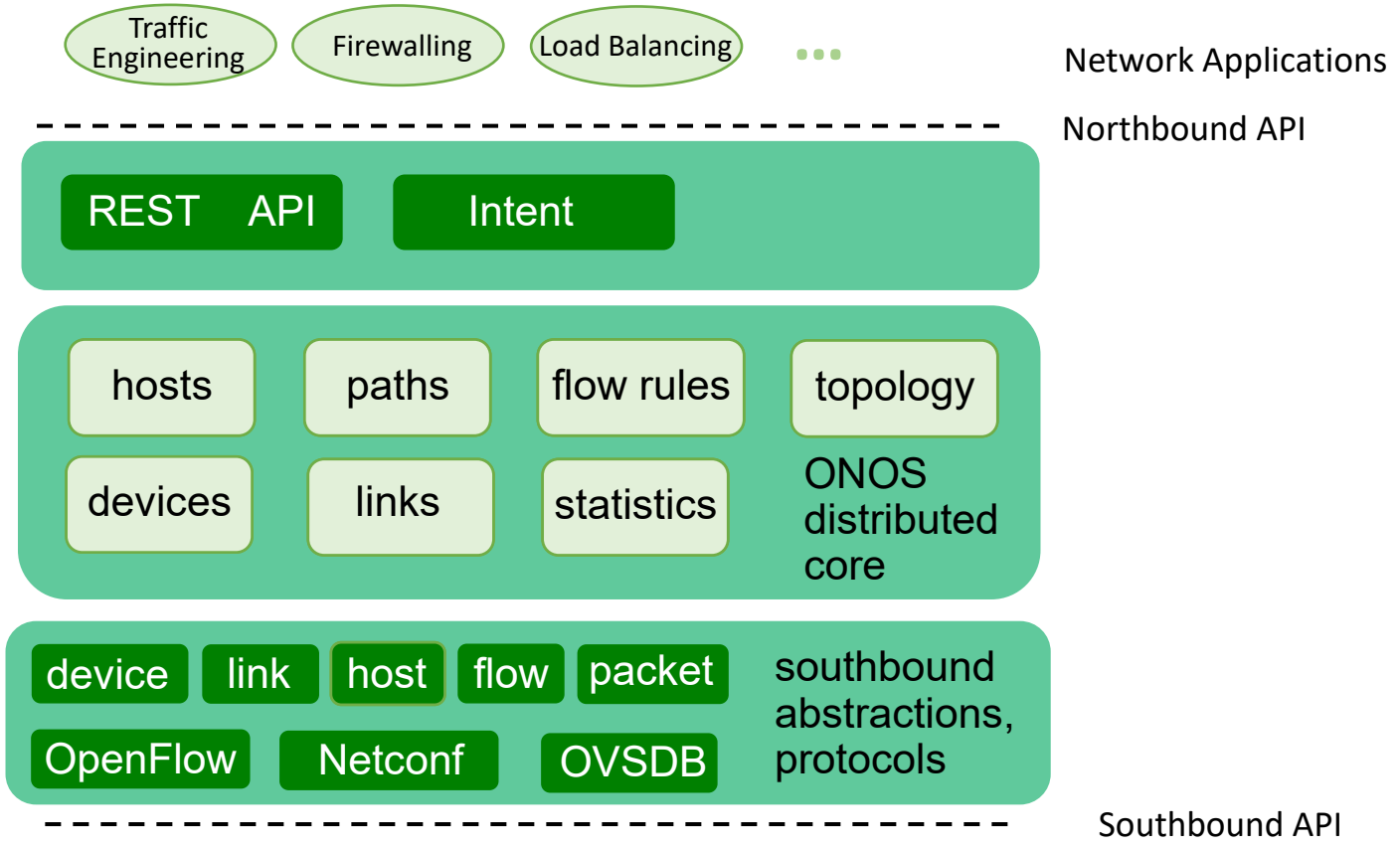
OpenDaylight (ODL) controller



Service Abstraction Layer:

- interconnects internal, external applications and services

ONOS controller



- control apps separate from controller
- intent framework: high-level specification of service: what rather than how
- considerable emphasis on distributed core: service reliability, replication performance scaling

SDN: sfide selezionate

- Hardening del piano di controllo: sistema distribuito *dependable*, scalabile nelle prestazioni e sicuro
  - robustezza ai guasti: sfruttare la teoria forte dei sistemi distribuiti affidabili per il piano di controllo
  - *dependability*, sicurezza: "incorporati" fin dal primo giorno?
- reti, protocolli che soddisfano i requisiti specifici di missione
  - es., tempo reale, ultra-affidabilità, ultra-sicurezza
- Estensione oltre un singolo AS
- L'SDN è fondamentale per le reti cellulari 5G

SDN e il futuro dei protocolli di rete tradizionali

- Tabelle di inoltro calcolate da SDN rispetto a quelle calcolate da router:
  - solo un esempio di calcolo logicamente centralizzato rispetto al calcolo protocollare
- si potrebbe immaginare un controllo della congestione calcolato da SDN:
  - il controllore imposta le velocità dei mittenti in base ai livelli di congestione segnalati dal router (al controllore)

Come evolverà l'implementazione delle funzionalità di rete (SDN rispetto ai protocolli)?