introducing COMPOSITIONAL NETWORK ARCHITECTURE

HOW THE INTERNET IS EVOLVING

Pamela Zave and Jennifer Rexford

Princeton University

Princeton, New Jersey, USA

IN I992

THE EXPLOSIVE GROWTH OF THE
WORLD-WIDE WEB BEGAN

AND IN 1993

THE LAST MAJOR CHANGE WAS MADE
TO THE “CLASSIC” INTERNET ARCHITECTURE

WHAT HAS HAPPENED SINCE 1993?

A FEW LITTLE THINGS:

- most of the world's

  . . . telecommunication infrastructure . . .

  . . . entertainment distribution . . .

  has moved to the Internet

- security threats are everywhere

- most networked devices are mobile

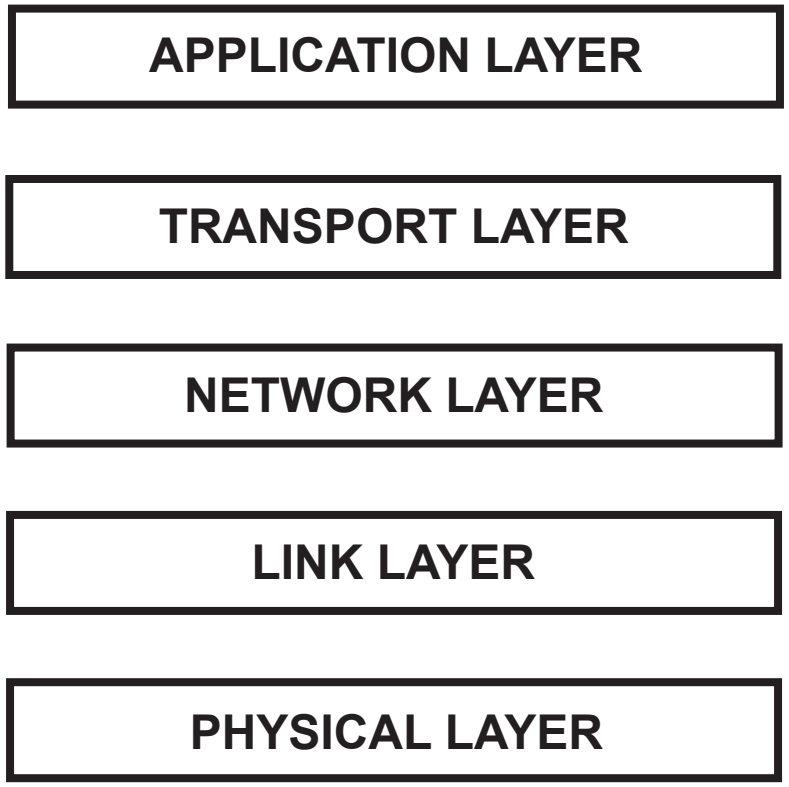
- many Internet services are hosted in a cloud

- the IPv4 address space is exhausted

- a cozy club of national telephone companies has been replaced by private enterprises in fierce competition

A MYSTERY:

THIS IS THE "CLASSIC" INTERNET ARCHITECTURE
(how experts describe the Internet)



IT HAS NOT CHANGED SINCE 1993, . . .

. . . YET THE INTERNET HAS MET ALL THESE NEW CHALLENGES (at least to some extent)

innovation within these layers,

without architectural change,

is *not enough!*

TO UNDERSTAND EVOLUTION, WE NEED TO LOOK AT THE BIG PICTURE

1 What is a NETWORK?

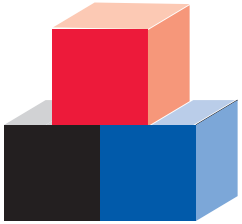
we need an abstract description that applies to many kinds of network → then we can view a network as a module or building block

2 The Internet ecosystem is a COMPOSITION of NETWORKS.

assembly *building blocks*

3 There are three COMPOSITION OPERATORS on NETWORKS.

ways of putting the blocks together

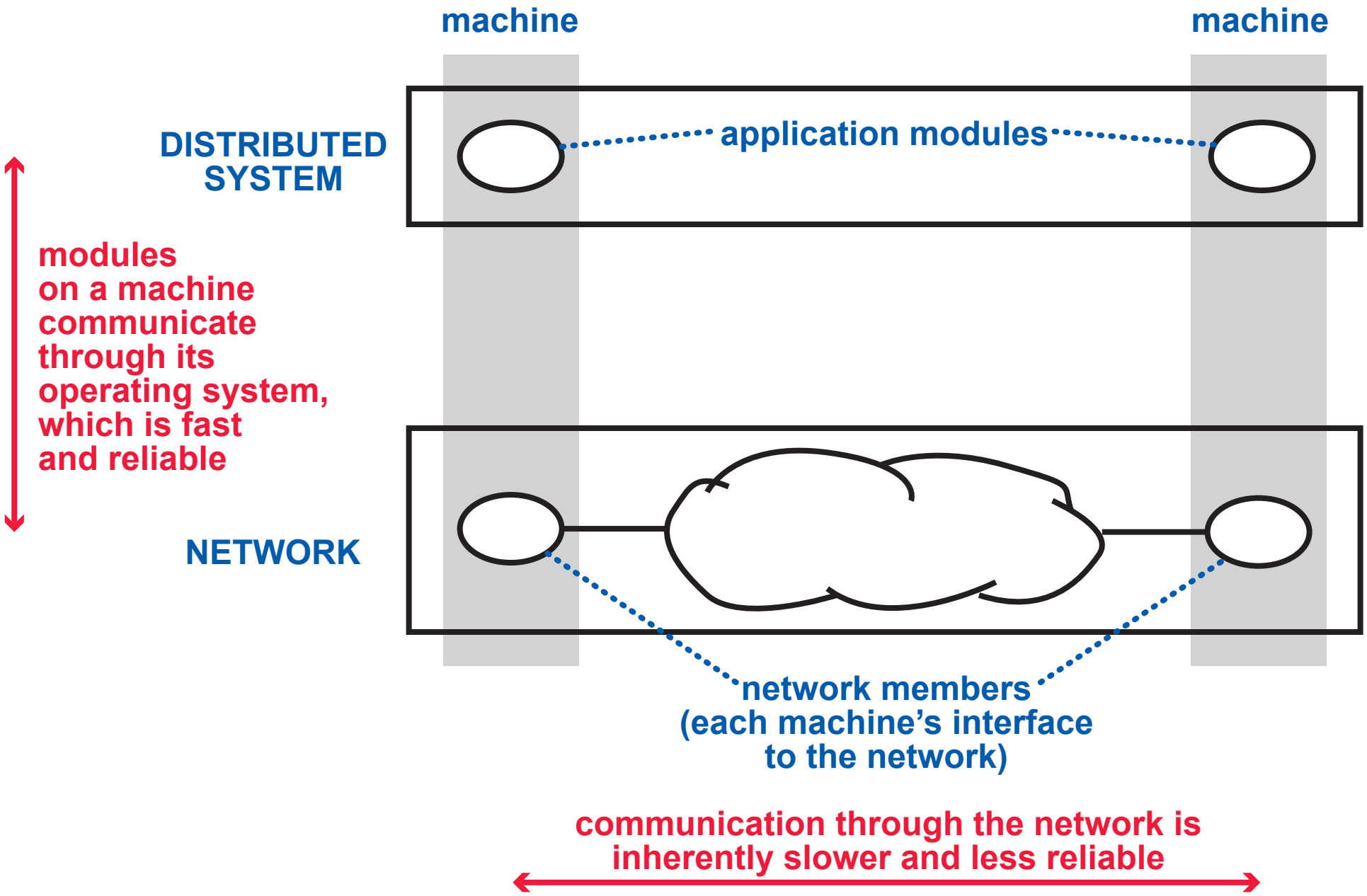


4 A need for EVOLUTION causes a PROBLEM;

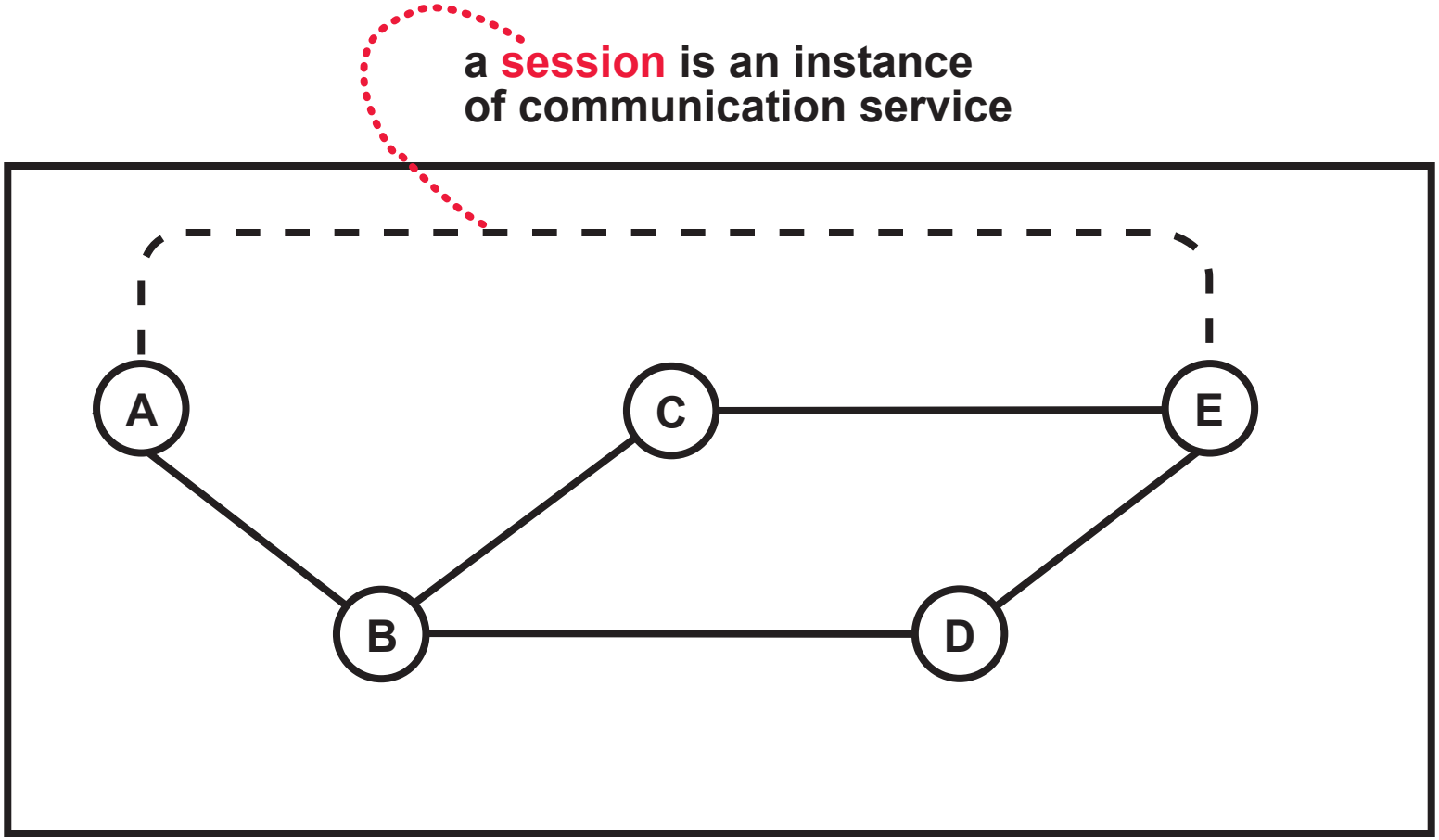
an ARCHITECTURAL PATTERN says "use this form of NETWORK COMPOSITION".

a problem-solving strategy, reusable in many different contexts

NETWORKS SUPPORT DISTRIBUTED SYSTEMS BY PROVIDING THEM WITH COMMUNICATION SERVICES



AN ABSTRACT NETWORK



AN ABSTRACT NETWORK

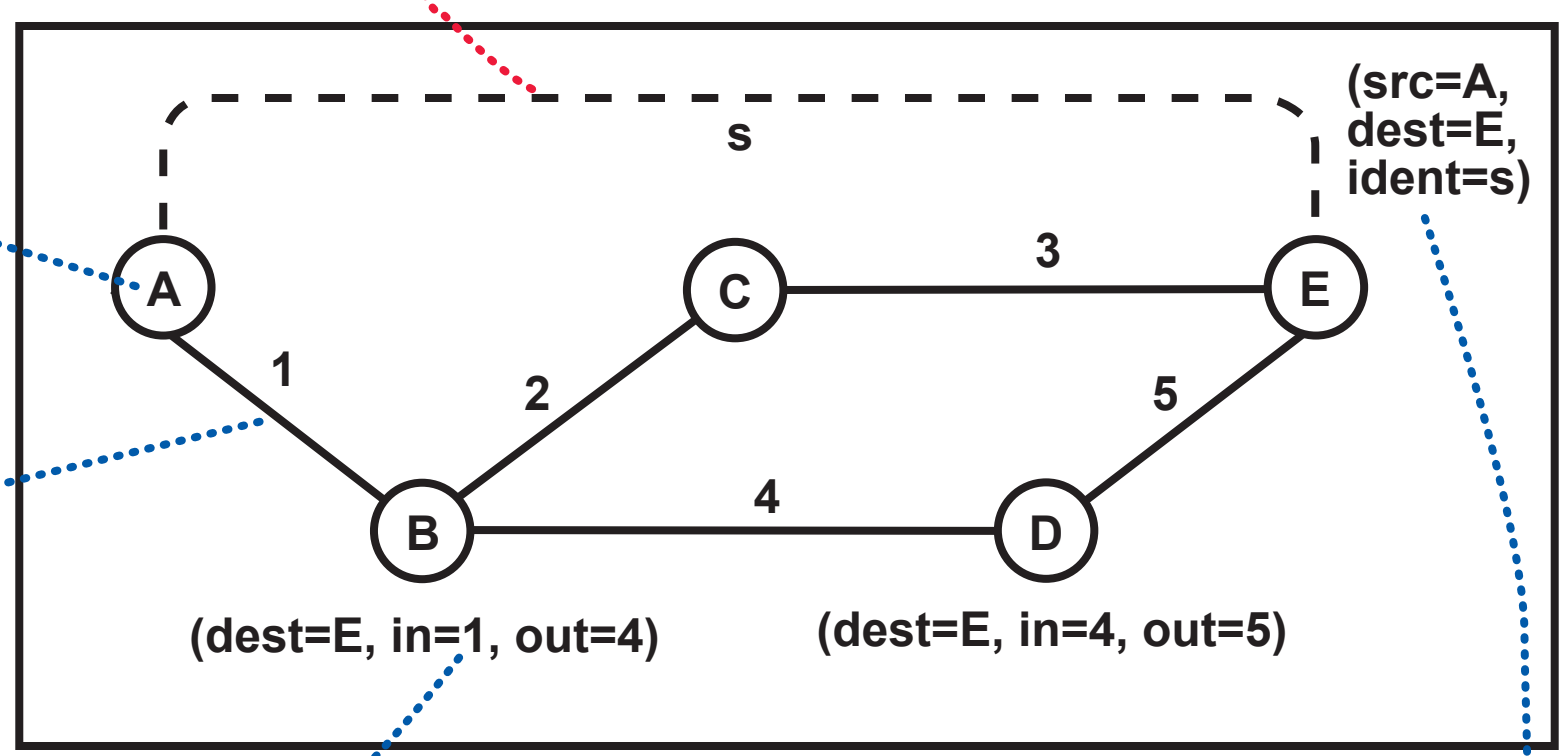
a session is an instance of communication service

members have names from a namespace

members are connected by links

routing chooses paths and populates forwarding tables, . . .

. . . which are used by the forwarding protocol



a network is

a self-contained

microcosm of

networking!

the service is implemented by a session protocol, . . .

. . . with session state in members

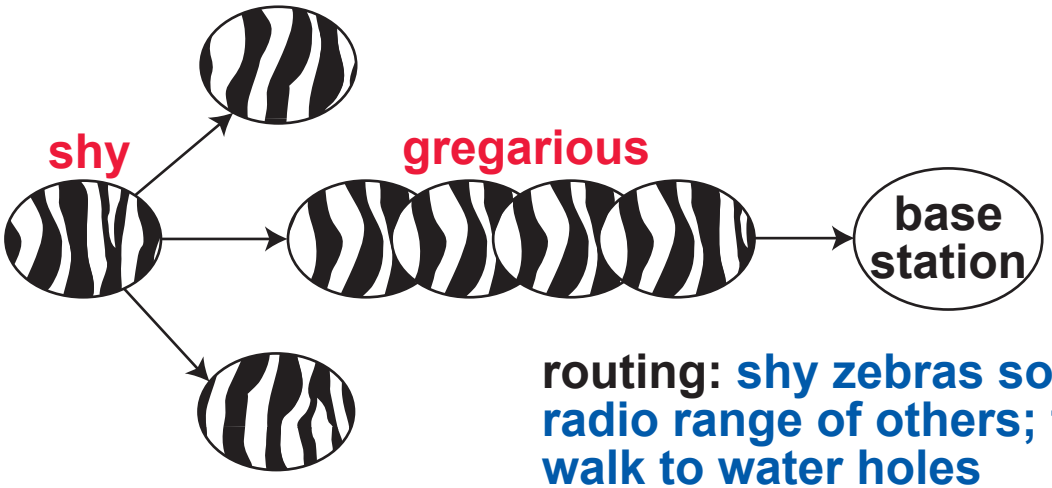
THERE ARE MANY KINDS OF NETWORK

Do you think that all networks have routing based on some graph algorithm?

ZEBRANET

members:
collars on zebras

links: radio



routing: shy zebras sometimes get within radio range of others; friendly zebras walk to water holes

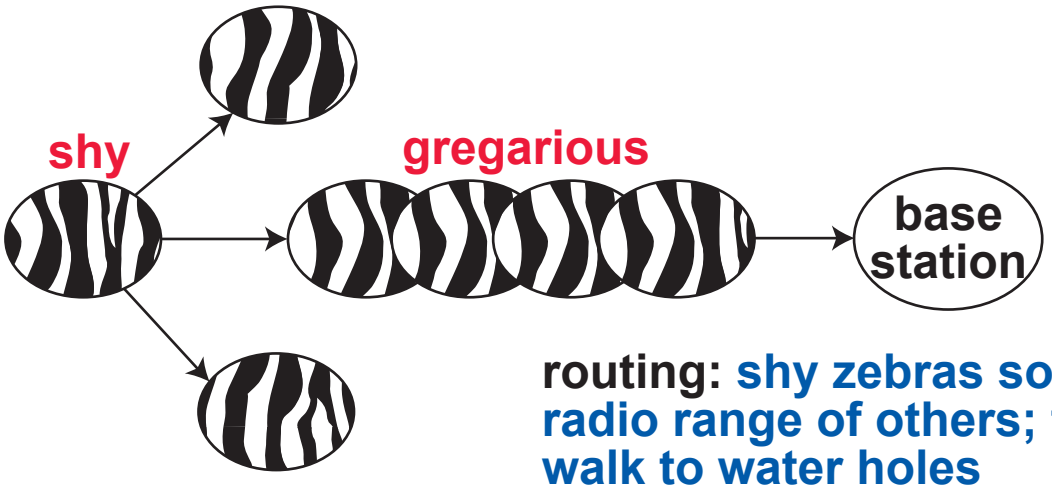
THERE ARE MANY KINDS OF NETWORK

Do you think that all networks have routing based on some graph algorithm?

ZEBRANET

members: collars on zebras

links: radio

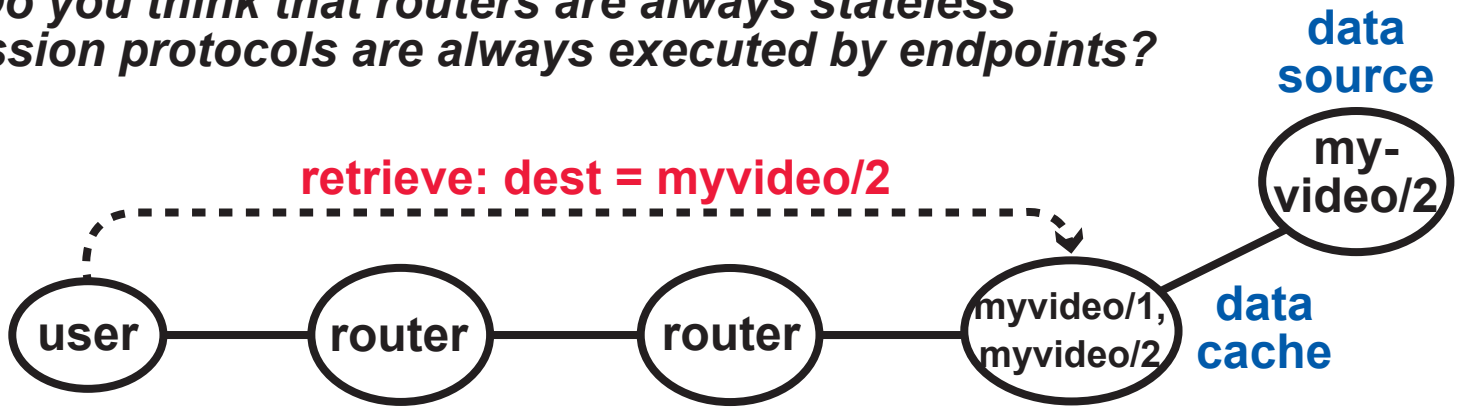


routing: shy zebras sometimes get within radio range of others; friendly zebras walk to water holes

Do you think that routers are always stateless and session protocols are always executed by endpoints?

NAMED DATA NETWORK

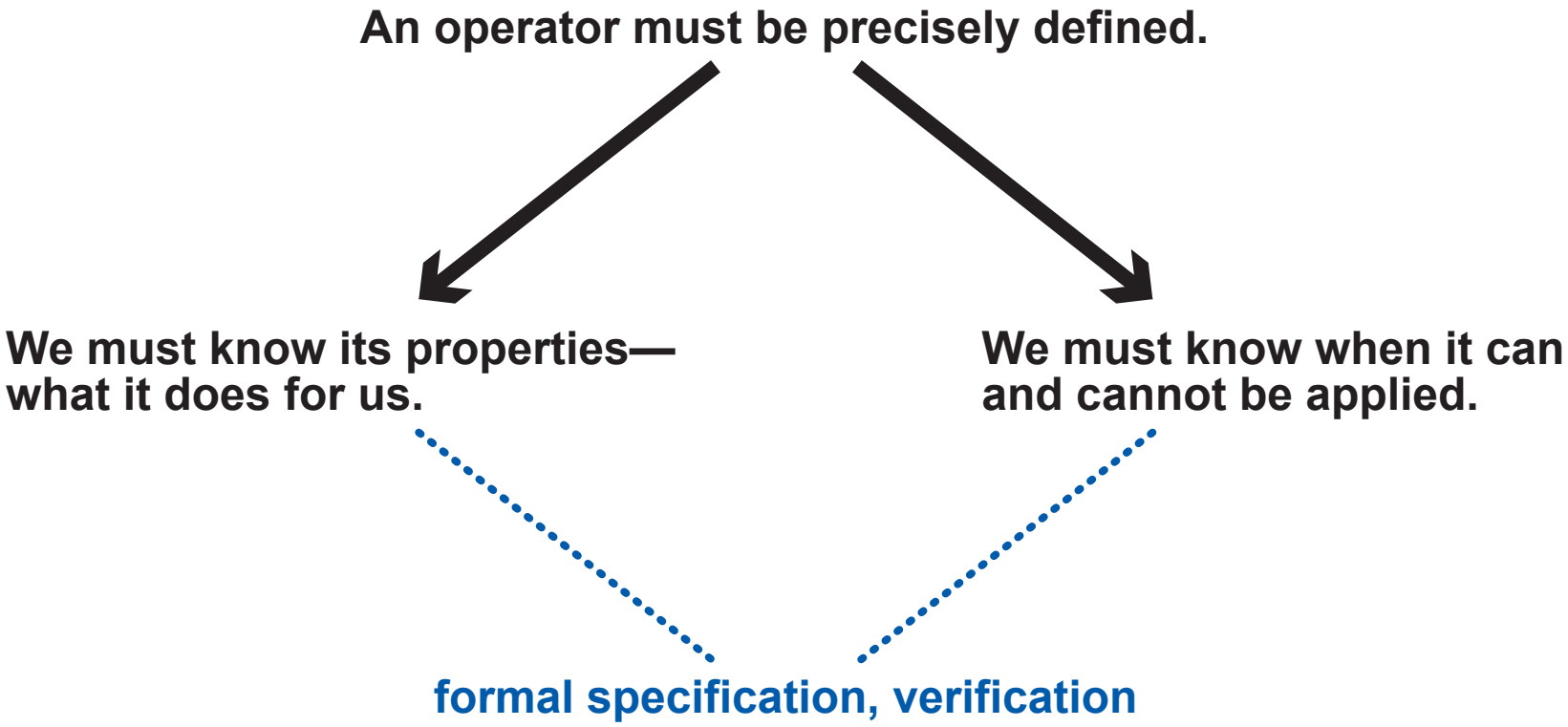
name: names a chunk of data, *and* is an anycast name for any member with the data



routing is normal—but routers also optimize sessions by caching requested data, bundling requests for the same data, etc.

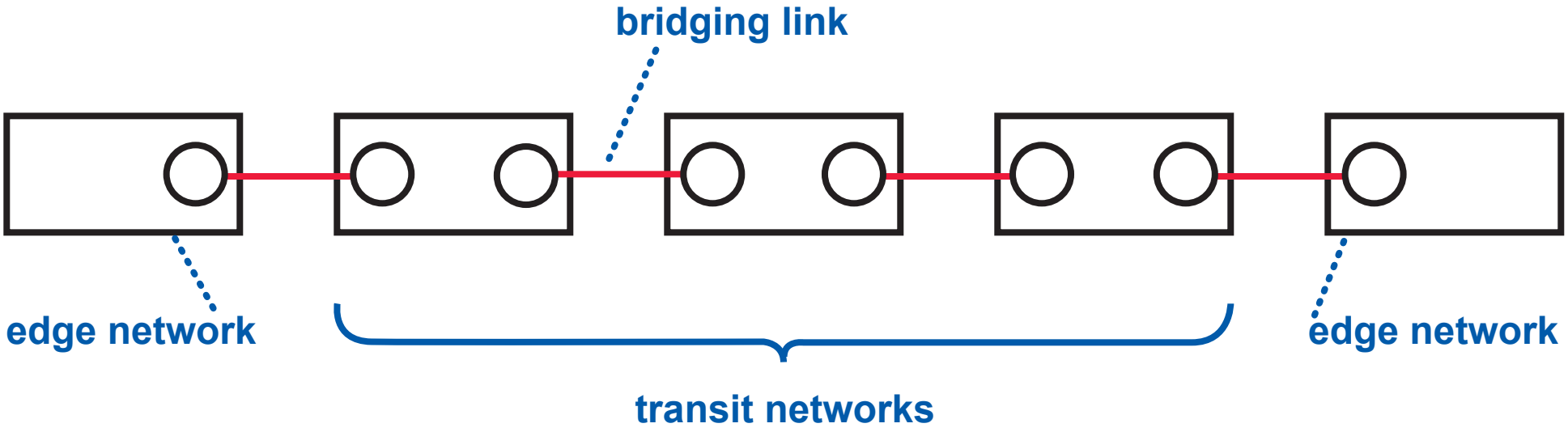
COMPOSITION OPERATORS ON NETWORKS

THE BIG PICTURE REQUIRES SOME NEW WAYS OF THINKING:



COMPOSITION OPERATORS ON NETWORKS 1

BRIDGING: The autonomous IP networks of the Internet are composed in this way.



benefit: bridging extends the reach of networks

limitation: the networks must be similar, because the same packets traverse multiple networks

COMPOSITION OPERATORS ON NETWORKS 2

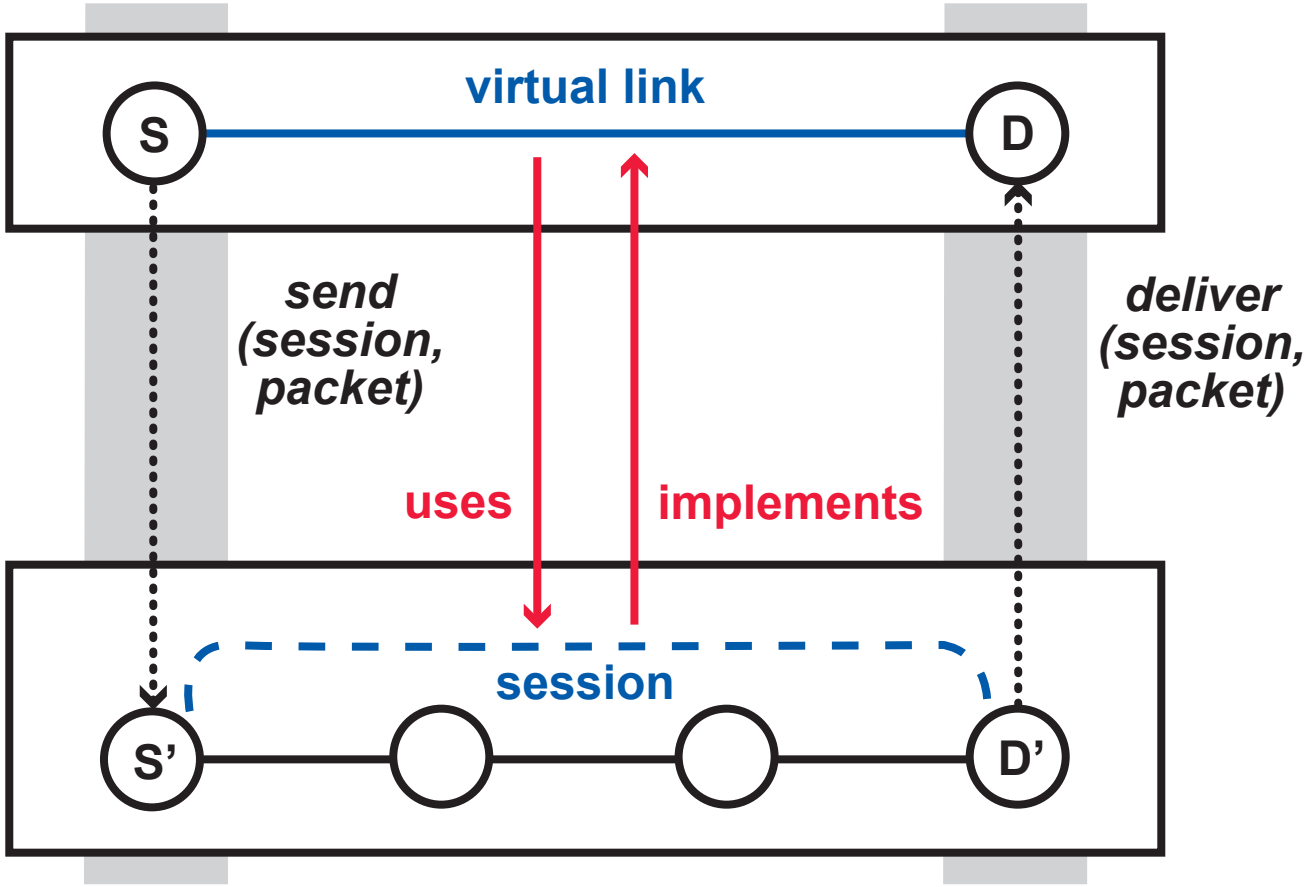
Any “layered architecture” has hierarchical dependencies.

Beyond that common ground, this definition is *specific* and *different*.

LAYERING:

A link of an overlay network . . .

is implemented by a session of an underlay network.

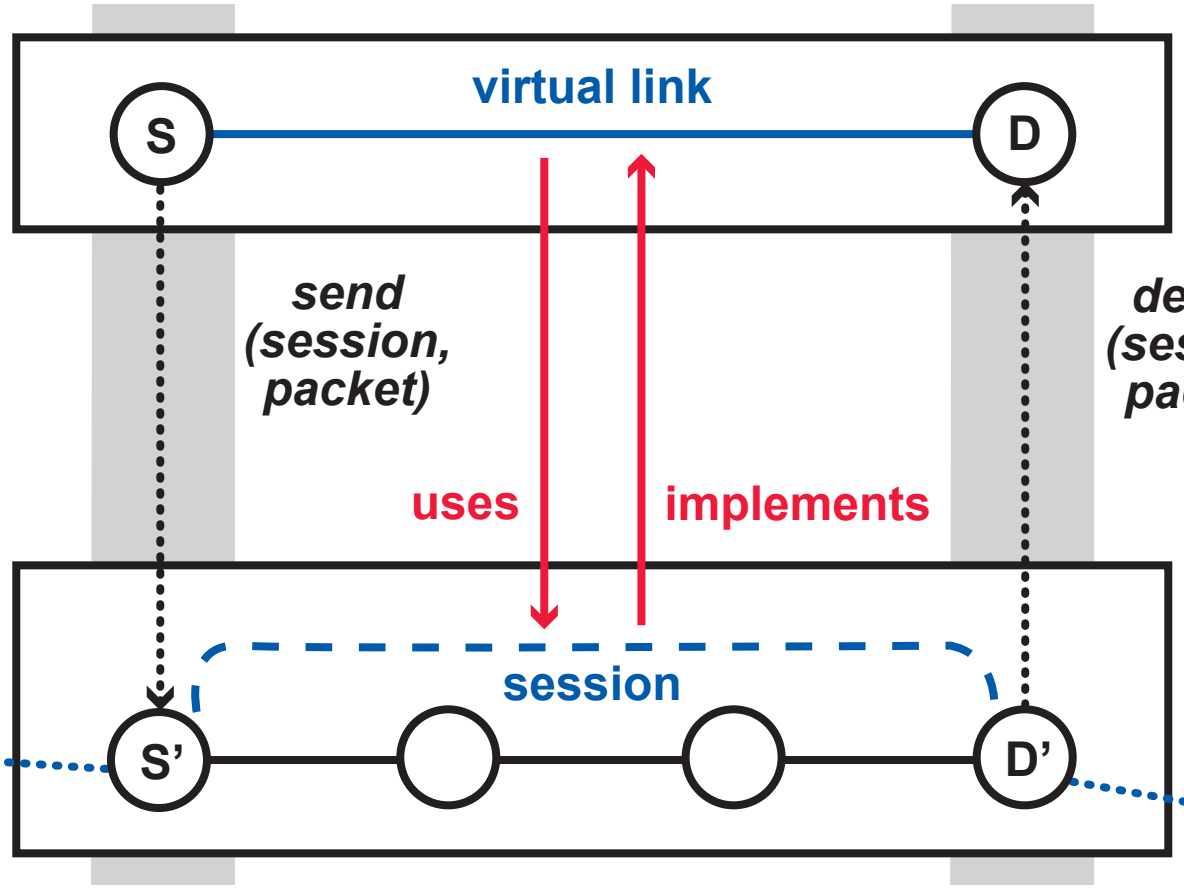


COMPOSITION OPERATORS ON NETWORKS 2

LAYERING:

A link of an overlay network . . .

is implemented by a session of an underlay network.



encapsulate packet in network and session header, . . . then transmit it

receive packet, . . . then decapsulate it (remove headers of this network)

Q:
Is this kind of layering for real?

A:
In the classic Internet architecture, a packet has 4 headers.
Today, it is easy to observe an Internet packet with 12.

THE BENEFITS OF LAYERING ARE MANY

overlay can have its own namespace

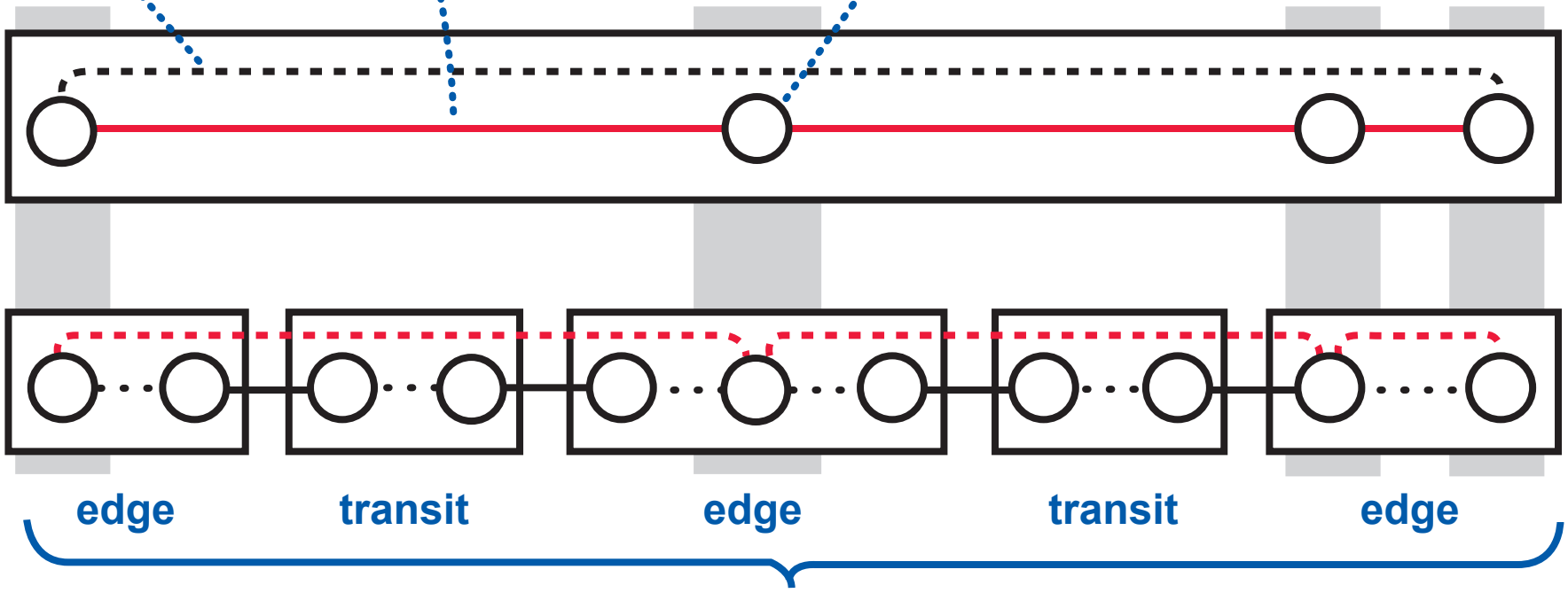
then it needs a directory to map overlay names to underlay names, e.g., DNS

overlay can have middleboxes to perform new functions (like security)

a virtual link can go anywhere

overlay can have new (or additional) session protocols

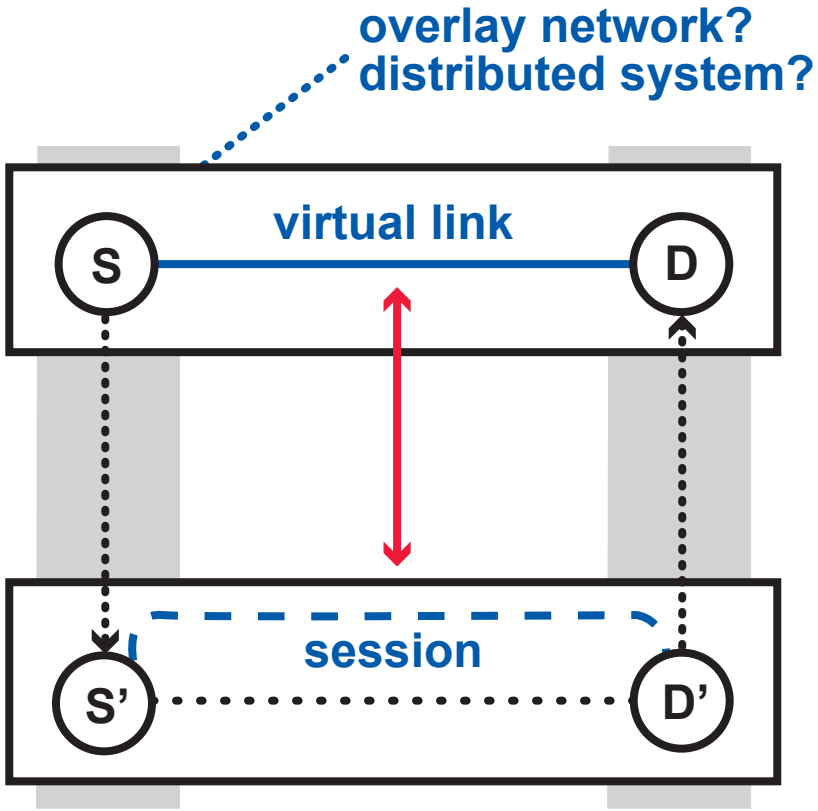
overlay can have its own routing algorithm, possibly to route session packets through the right middleboxes



bridged networks of the base Internet

THE LIMITATIONS OF LAYERING ARE FEW

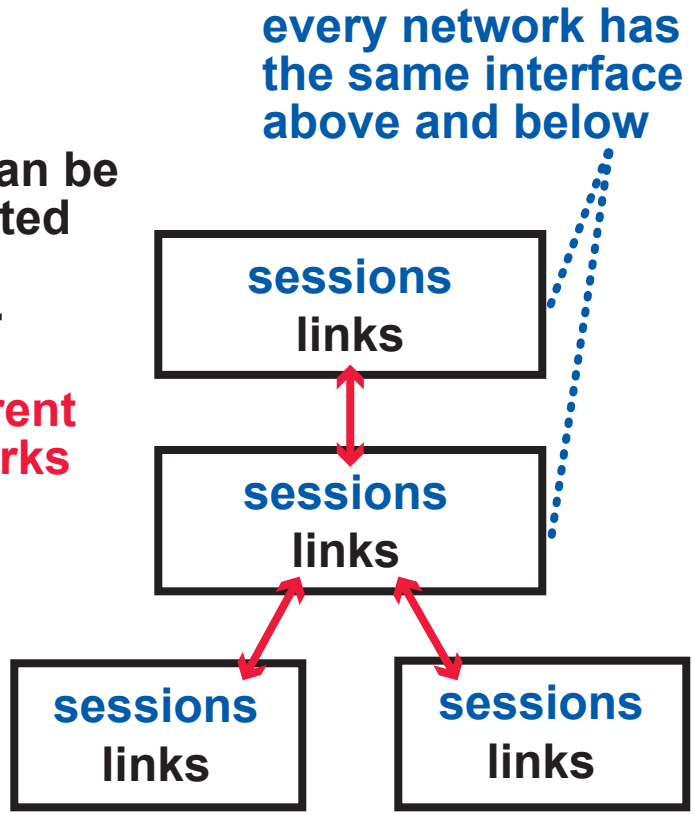
THE LAYERING INTERFACE IS THE SAME AS THE USER INTERFACE!



- anyone can add an overlay network to the Internet ecosystem
- this is how the Internet has evolved dramatically without changes to IP standards

NETWORKS CAN BE COMPOSED AS FREELY AS LEGO BLOCKS!

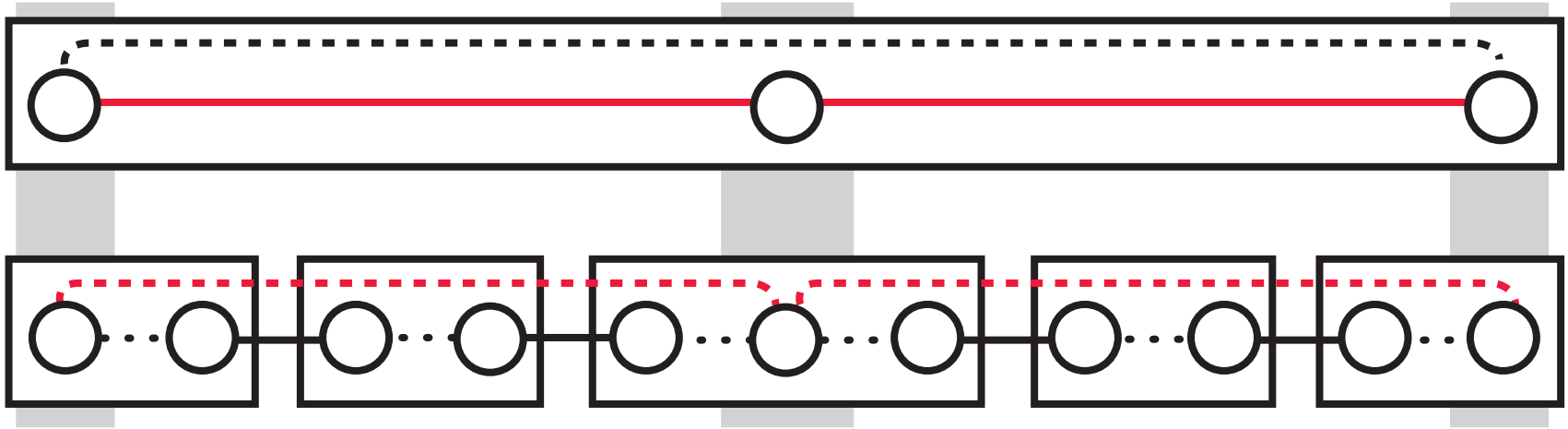
any link can be implemented by any session— no matter how different the networks are



- this does not work in the classic Internet architecture because each layer has different functions, different interfaces

THE LIMITATION OF LAYERING

machines with overlay members
have some right to
use the services . . .
. . . or maybe they have
just downloaded the
“client” software



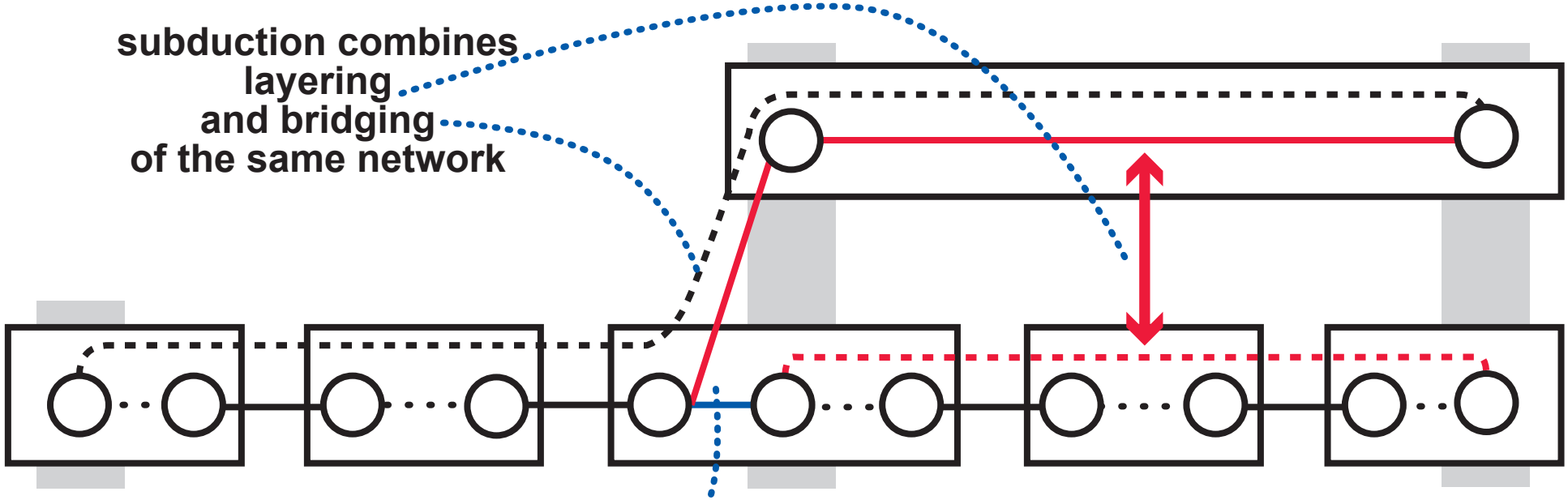
but what if a member of the overlay
wants to communicate with a machine
that does not have a member . . .

. . . while using the services of the overlay?

COMPOSITION OPERATORS ON NETWORKS 3

SUBDUCTION: With subduction, machines with a service can communicate with machines that don't have the service.

Subduction is the enabler of innovation in the Internet!



subduction requires special shared links . . .

. . . and defines how the two packet paths are coordinated

formalization shows that subduction preserves the properties of both bridging and layering

THAT'S ENOUGH THEORY!

NOW LET'S LOOK AT REAL EXAMPLES OF EVOLUTION

SUMMARY OF AN EXAMPLE:

evolution
. . . to meet a new need

poses a problem
. . . because the need cannot be met within the existing architecture

a pattern solves the problem
. . . with a new architecture, realized by composing networks!

BASE INTERNET 1: BEFORE THERE WAS AN INTERNET

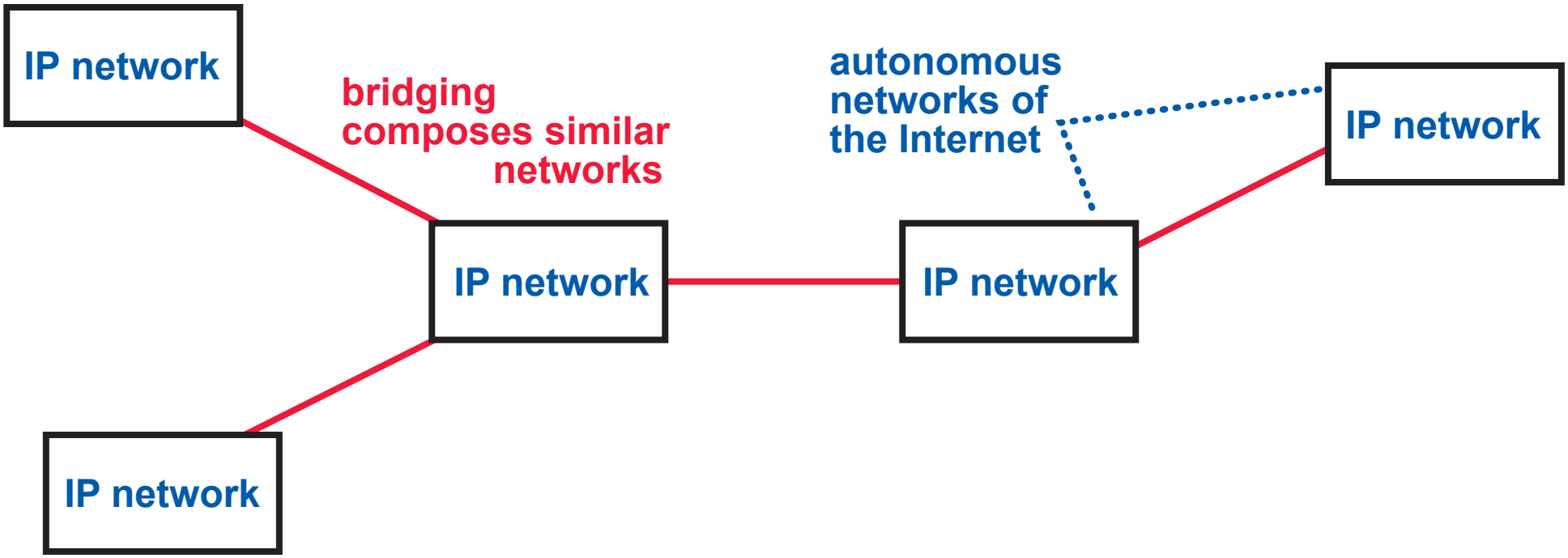
There are local networks all over the world.

radio networks (WiFi, cellular, mobile ad hoc)	wired networks (Ethernets, buses, cables)	optical and microwave networks

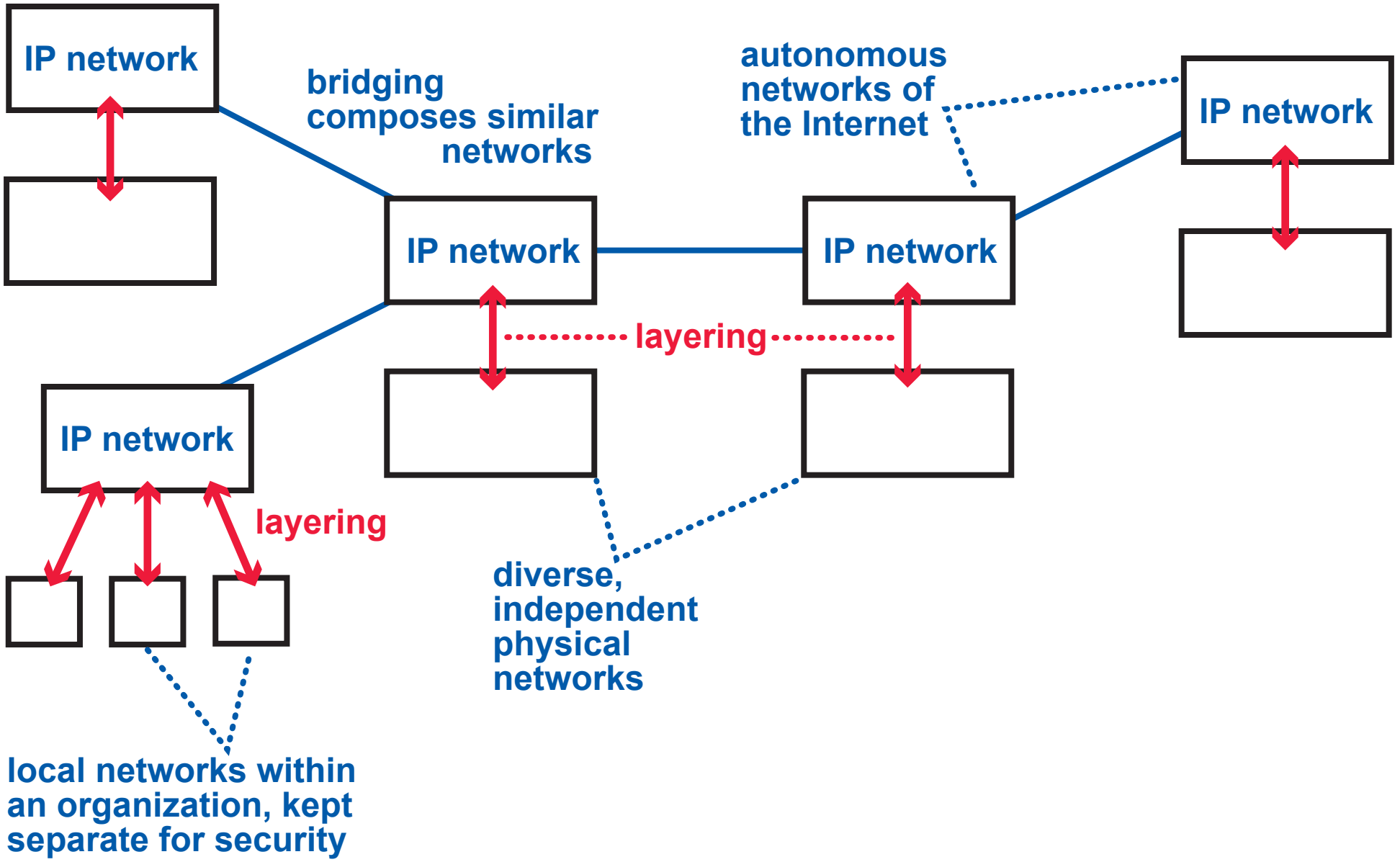
There is a need for them to communicate, but how?

There are many different kinds, with different protocols.

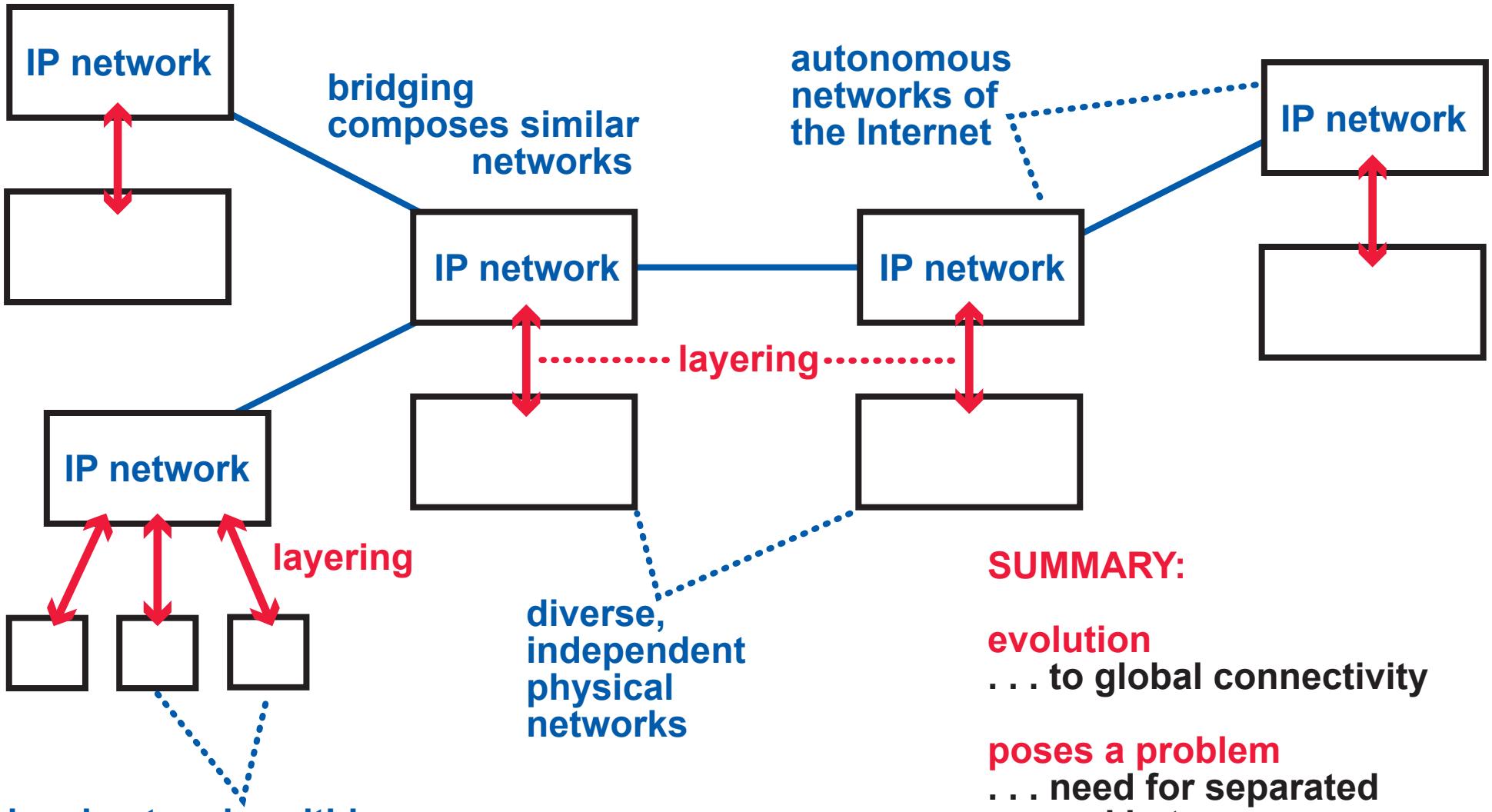
BASE INTERNET 2: THE INTERNET PROTOCOL IS GLOBAL



BASE INTERNET 2: THE INTERNET PROTOCOL IS GLOBAL



BASE INTERNET 2: THE INTERNET PROTOCOL IS GLOBAL



SUMMARY:

evolution
. . . to global connectivity

poses a problem
. . . need for separated and heterogeneous networks to interoperate

pattern solves the problem
. . . with bridged and layered networks

MOBILITY 1: CELLPHONES, LAPTOPS, AND THE INTERNET OF THINGS

TODAY, MOST NETWORKED DEVICES ARE MOBILE . . .

. . . BUT THE BASE INTERNET IS NOT DESIGNED FOR MOBILITY

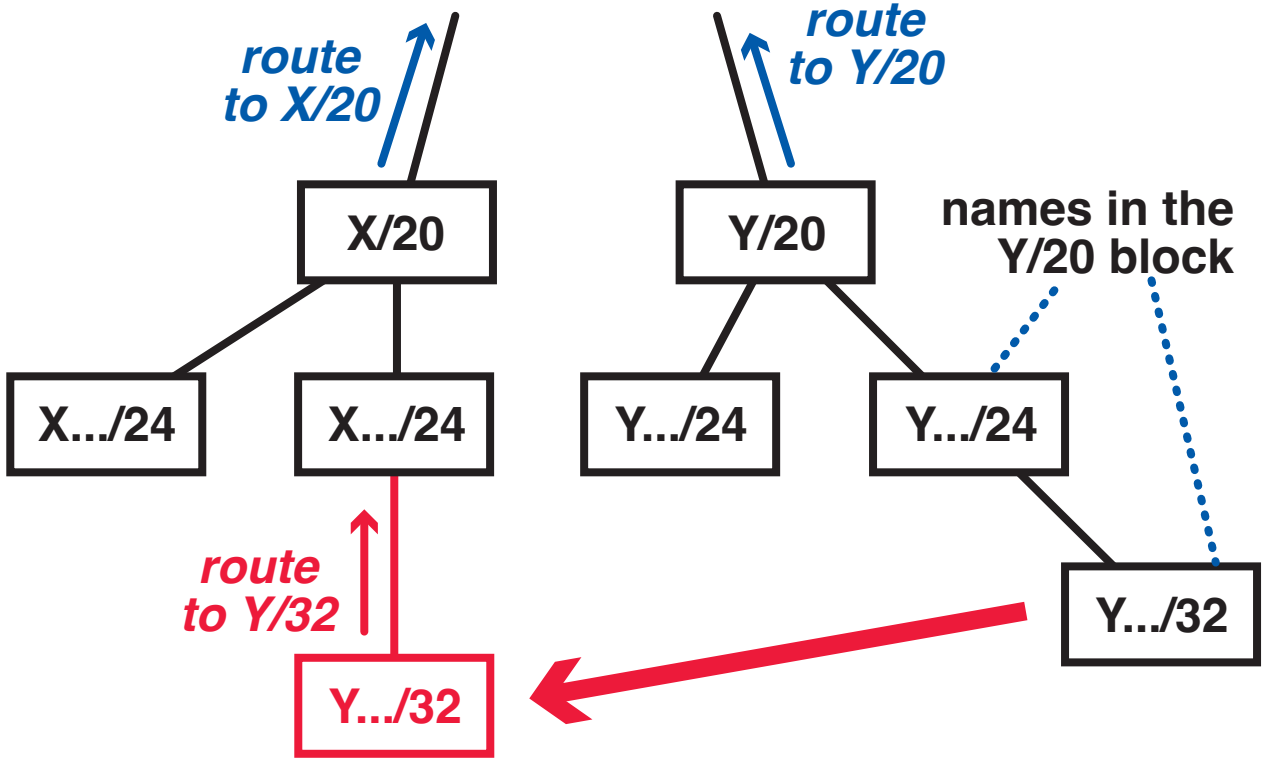
in the strongest sense: session continues as an endpoint moves

routing works at Internet scale because names are hierarchical and location-dependent, . . .

. . . and routes to large blocks are advertised

for session preservation, a member's name must move with it

but this means that mobile names must be advertised, *frequently,* as blocks of size 1!

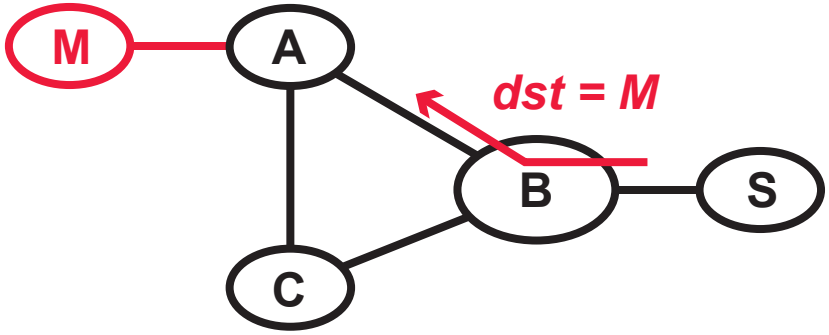


Boeing tried this once, causing howls of protest from network operators everywhere

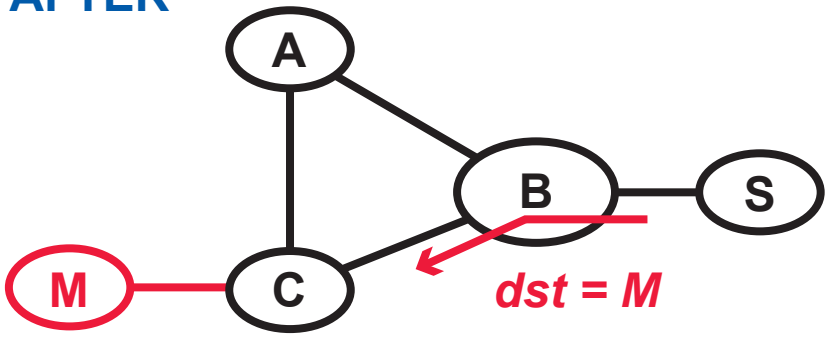
MOBILITY 2: TWO PATTERNS FOR MOBILITY

DYNAMIC ROUTING MOBILITY

BEFORE



AFTER



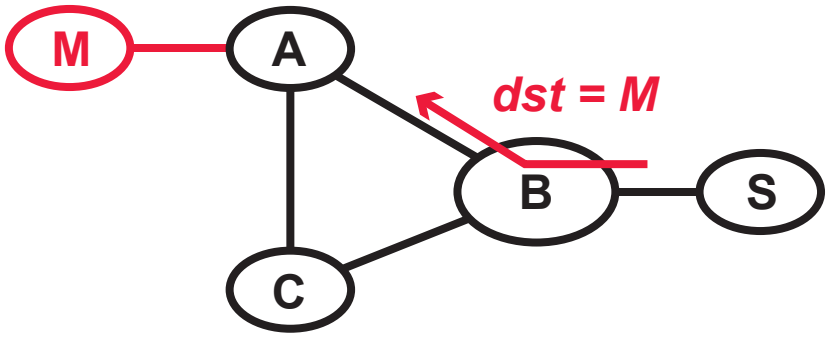
routing is updated dynamically to reach the mobile member in its new location

this is what *cannot* be done in the base Internet

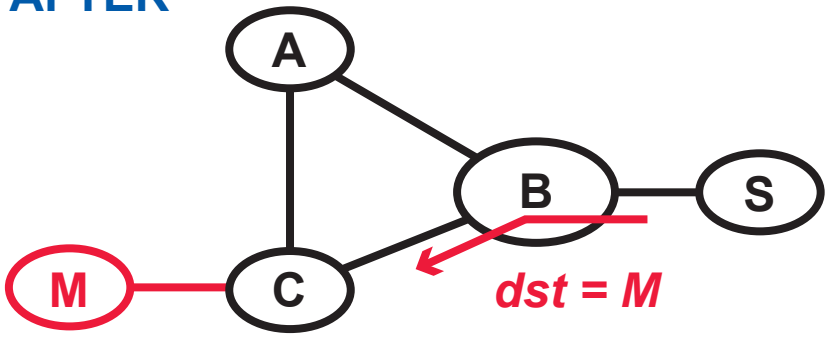
MOBILITY 2: TWO PATTERNS FOR MOBILITY

DYNAMIC ROUTING MOBILITY

BEFORE



AFTER

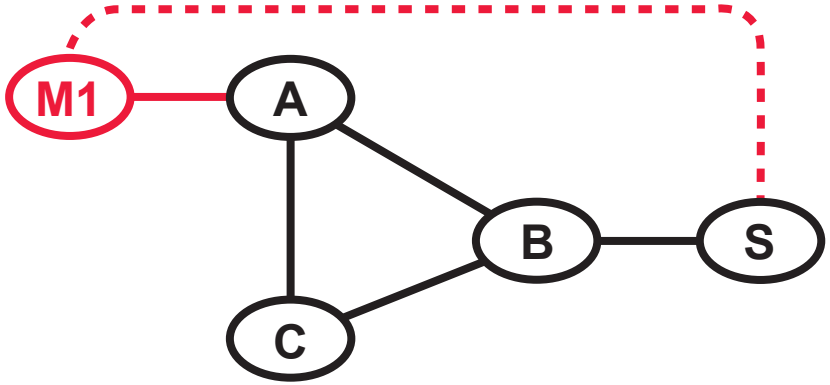


routing is updated dynamically to reach the mobile member in its new location

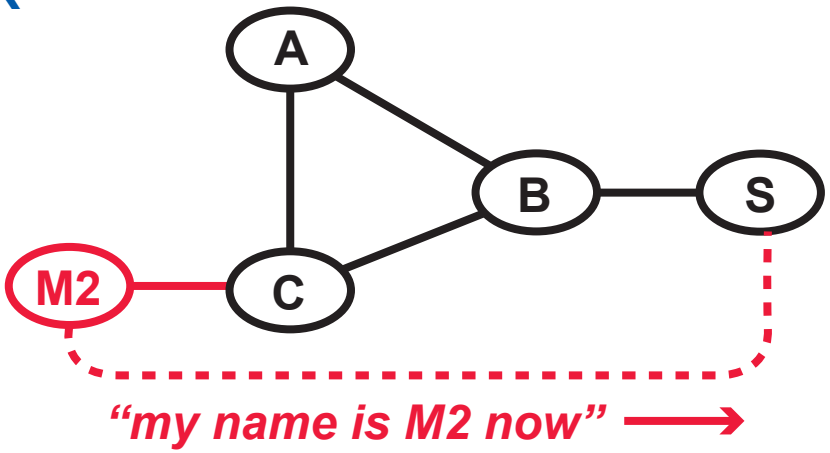
this is what *cannot* be done in the base Internet

SESSION-LOCATION MOBILITY

BEFORE



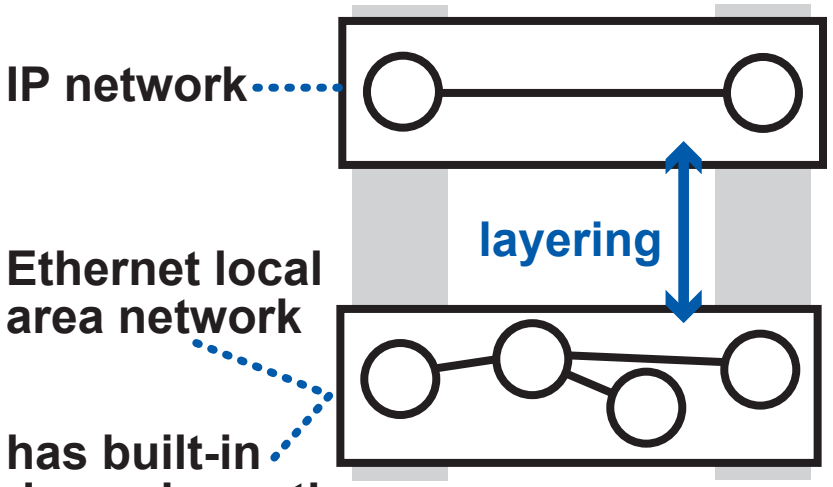
AFTER



in a location-dependent namespace, mobile member changes name as it moves . . . and uses the session protocol to tell the other endpoint its new name

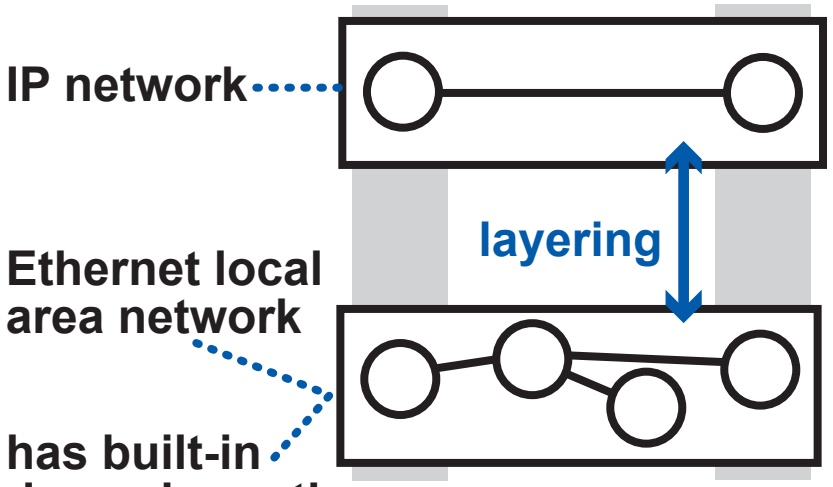
MOBILITY 3: DYNAMIC ROUTING MOBILITY

IN AN EDGE NETWORK OF THE INTERNET



MOBILITY 3: DYNAMIC ROUTING MOBILITY

IN AN EDGE NETWORK OF THE INTERNET

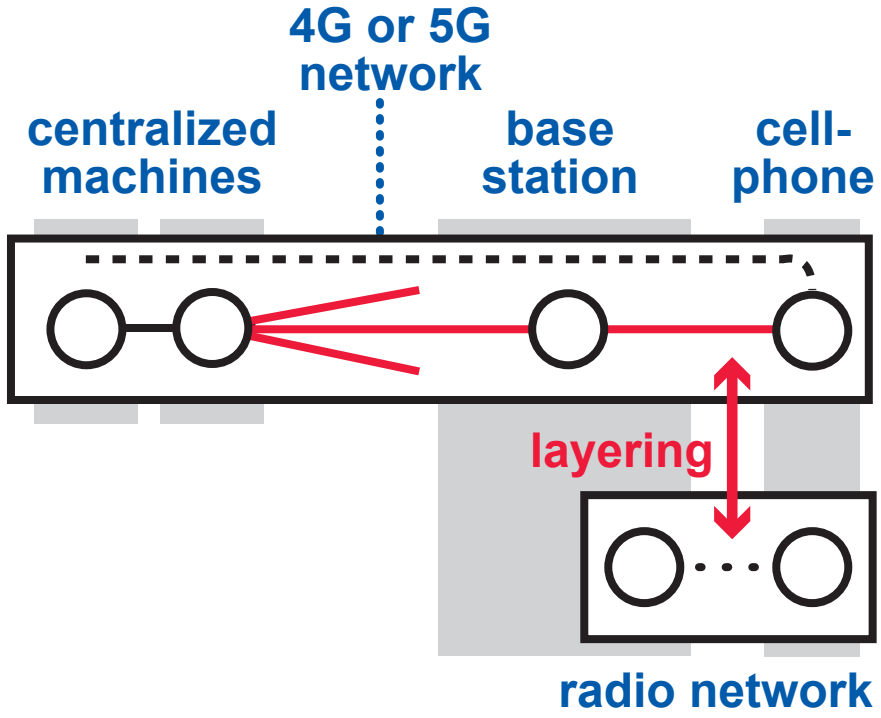


. . . which works for a small-scale network

IN A 4G OR 5G NETWORK

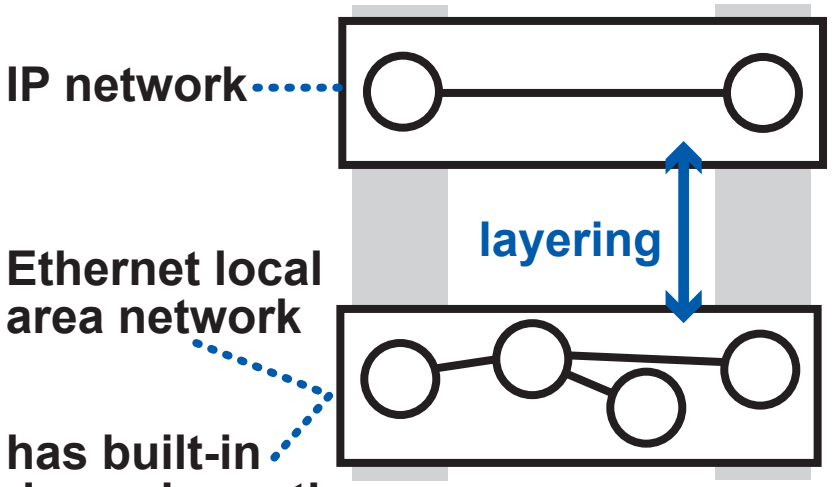
centralized machines maintain virtual links to every base station, and maintain a current route to each cellphone

this is expensive, but not compared to running a cellular network!



MOBILITY 3: DYNAMIC ROUTING MOBILITY

IN AN EDGE NETWORK OF THE INTERNET

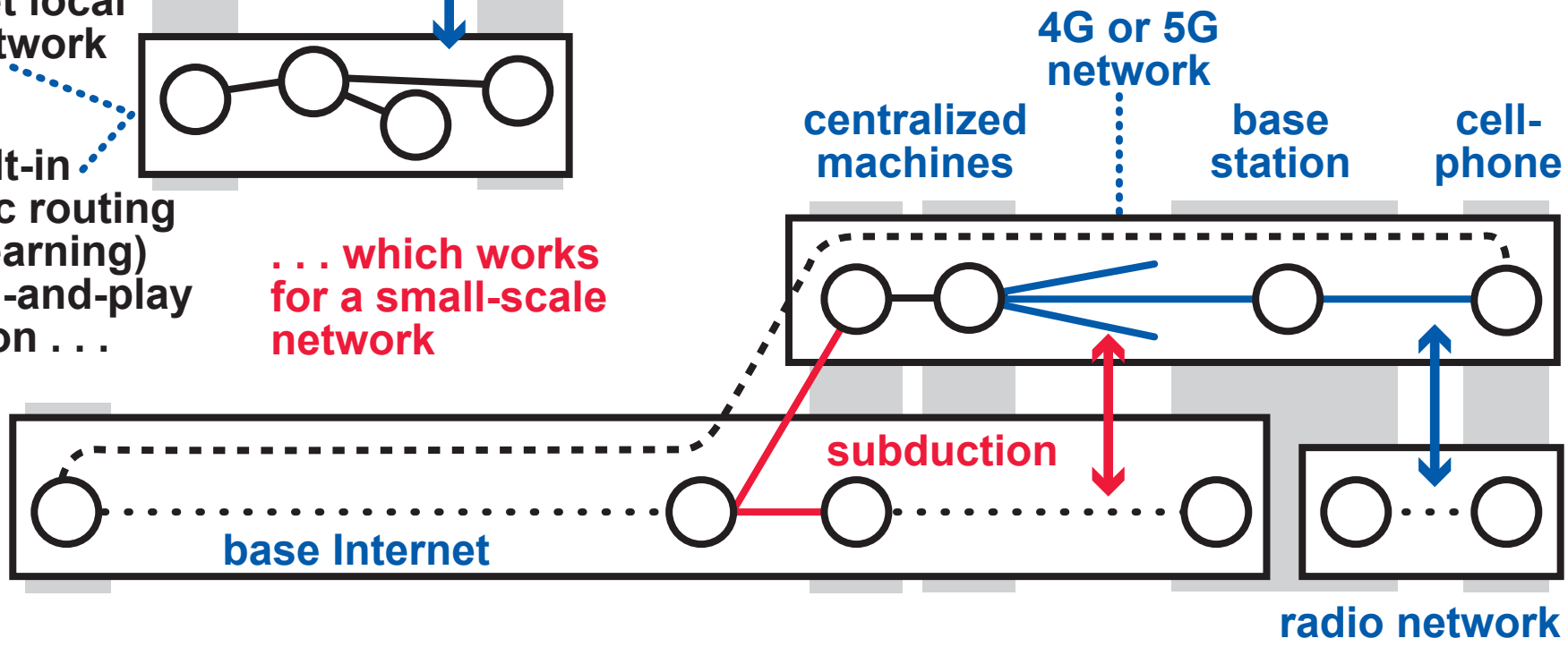


. . . which works for a small-scale network

IN A 4G OR 5G NETWORK

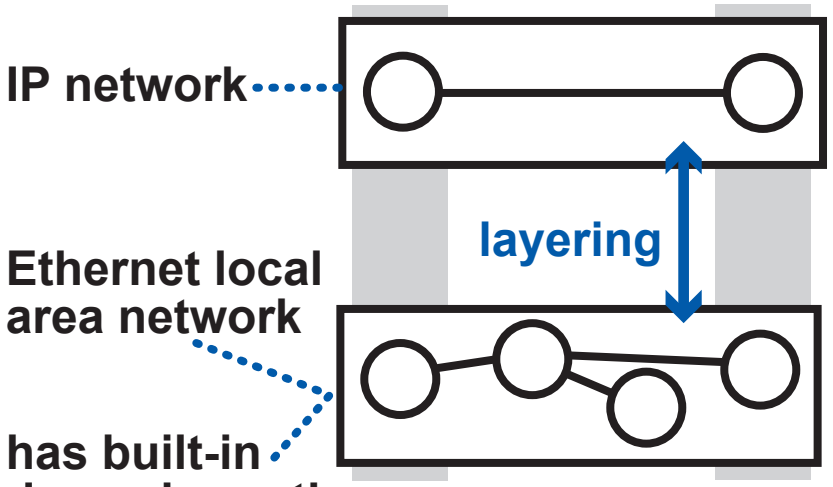
centralized machines maintain virtual links to every base station, and maintain a current route to each cellphone

this is expensive, but not compared to running a cellular network!



MOBILITY 3: DYNAMIC ROUTING MOBILITY

IN AN EDGE NETWORK OF THE INTERNET

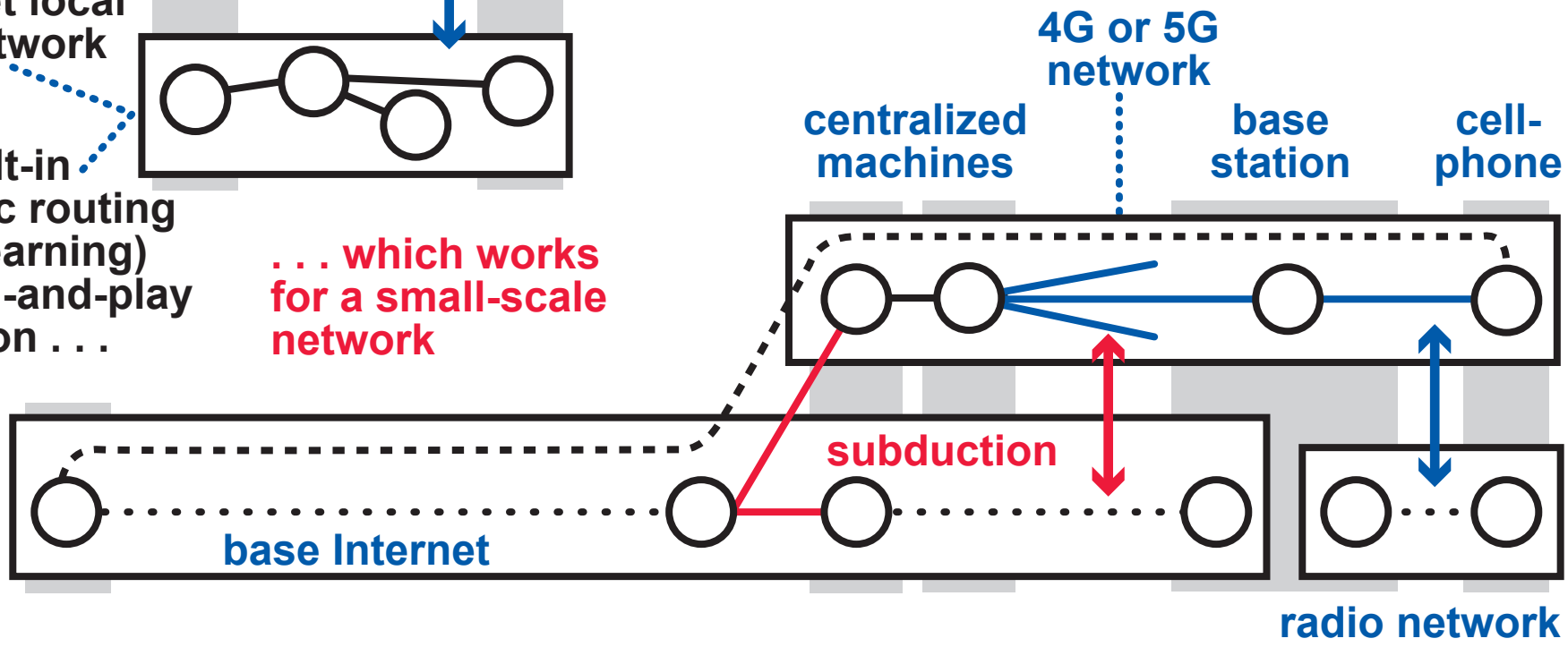


. . . which works for a small-scale network

IN A 4G OR 5G NETWORK

centralized machines maintain virtual links to every base station, and maintain a current route to each cellphone

this is expensive, but not compared to running a cellular network!



evolution . . . to add an enhanced service
poses a problem . . . that the classic Internet architecture cannot solve
DRM solves the problem . . . with dynamic routing in smaller networks, composing them with Internet by layering or subduction

MOBILITY 4: SESSION-LOCATION MOBILITY

CHEAP IN THE BASE INTERNET, . . .

. . . so there have been many research proposals . . .

BUT THERE ARE ALSO PROBLEMS

- need new session protocols

  QUIC and Mobile IP have SLM

- security risks (spoofing, directory update by user members)

  modern cryptography removes these

- how does a member connect to another member with no persistent name?

  QUIC is client/server, only clients are mobile

MOBILITY 4: SESSION-LOCATION MOBILITY

CHEAP IN THE BASE INTERNET, . . .

. . . so there have been many research proposals . . .

BUT THERE ARE ALSO PROBLEMS

- **need new session protocols**

  QUIC and Mobile IP have SLM

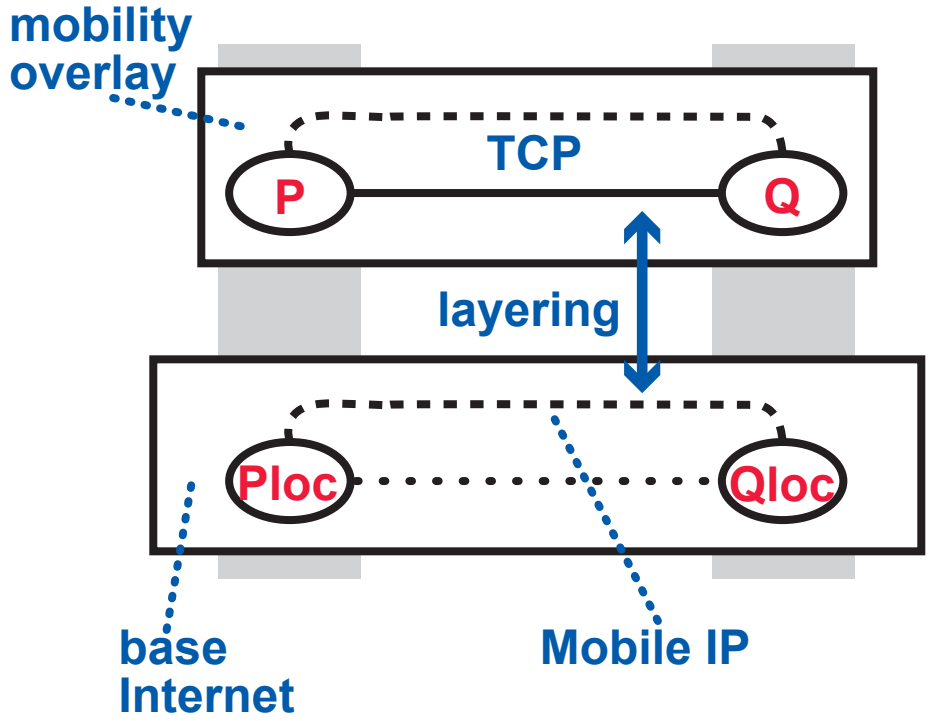
- **security risks (spoofing, directory update by user members)**

  modern cryptography removes these

- **how does a member connect to another member with no persistent name?**

  QUIC is client/server, only clients are mobile

session is set up in overlay using P and Q, which are persistent names for mobile machines



if only one endpoint is mobile? overlay contains one endpoint, subduction is used

MOBILITY 4: SESSION-LOCATION MOBILITY

CHEAP IN THE BASE INTERNET, . . .

. . . so there have been many research proposals . . .

BUT THERE ARE ALSO PROBLEMS

- **need new session protocols**

  QUIC and Mobile IP have SLM

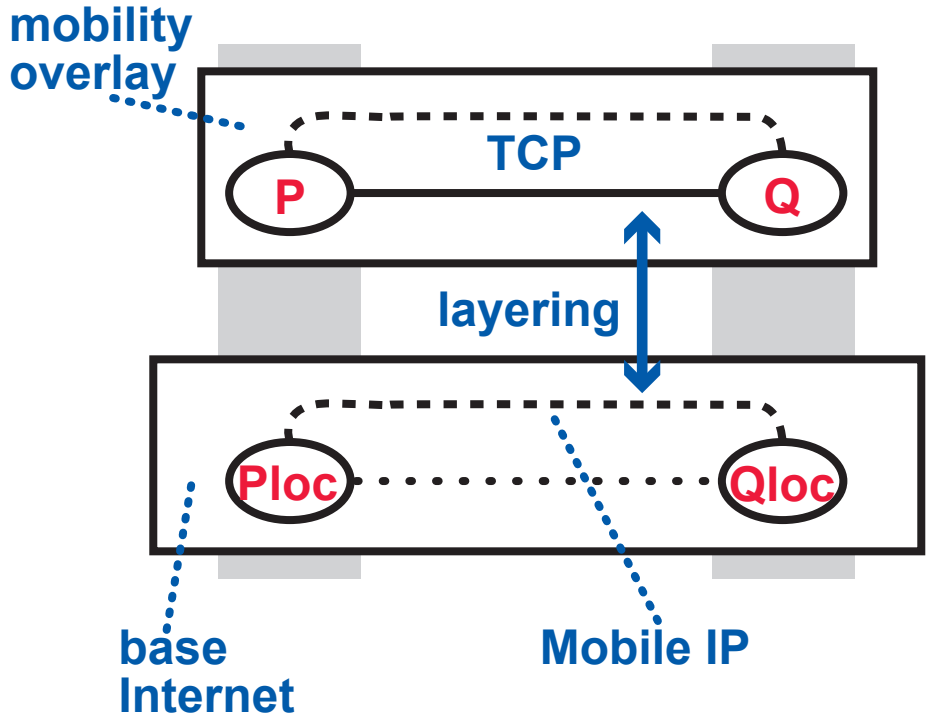
- **security risks (spoofing, directory update by user members)**

  modern cryptography removes these

- **how does a member connect to another member with no persistent name?**

  QUIC is client/server, only clients are mobile

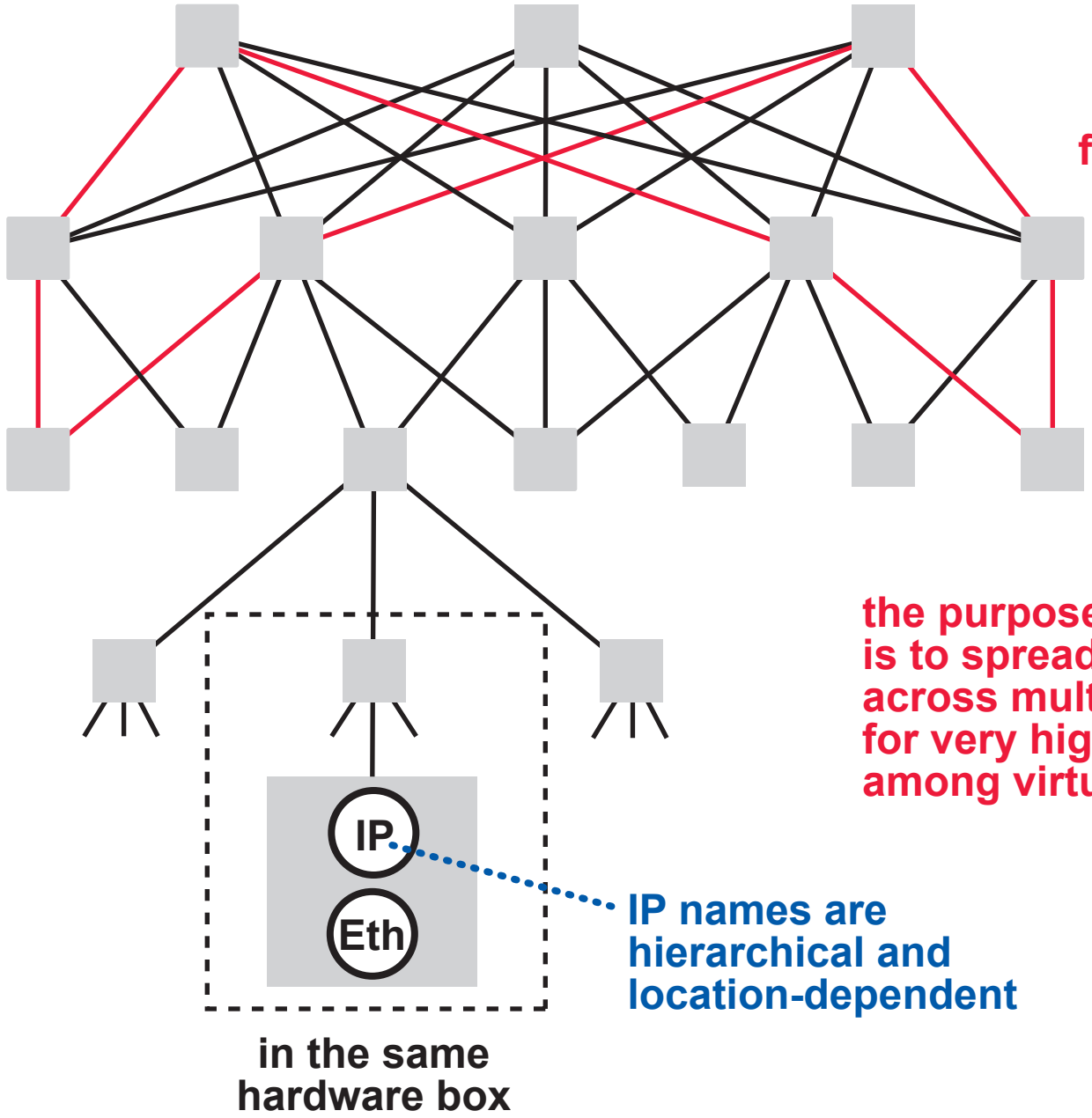
session is set up in overlay using P and Q, which are persistent names for mobile machines



if only one endpoint is mobile? overlay contains one endpoint, subduction is used

evolution . . . to add mobility
poses a problem . . . that members change Internet names as they move
SLM solves the problem . . . with overlays with location-independent names, composing them with Internet by layering or subduction

CLOUD COMPUTING 1: DATA-CENTER NETWORK



a data-center network has a fat-tree topology with up to 1 BILLION virtual machines

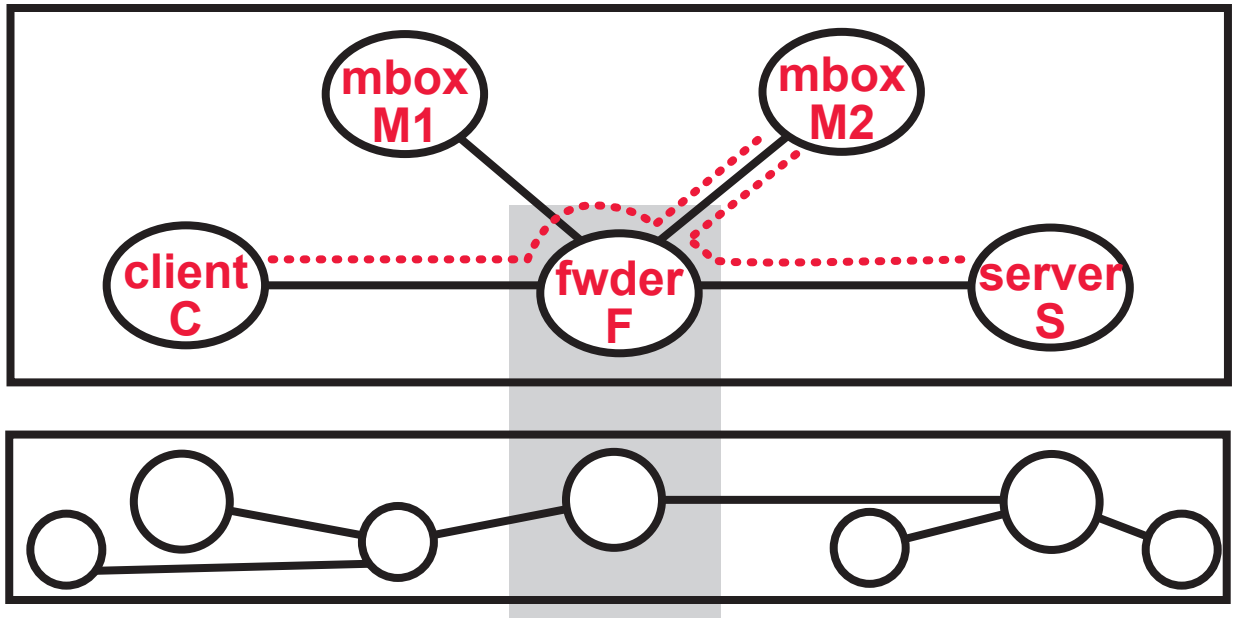
the purpose of IP routing is to spread traffic across multiple long paths, for very high throughput among virtual machines

CLOUD COMPUTING 2: ENTERPRISE NETWORK

IP network . . .

. . . has naming, topology, and routing designed for enterprise . . .

. . . and is layered on an Ethernet network that implements its links

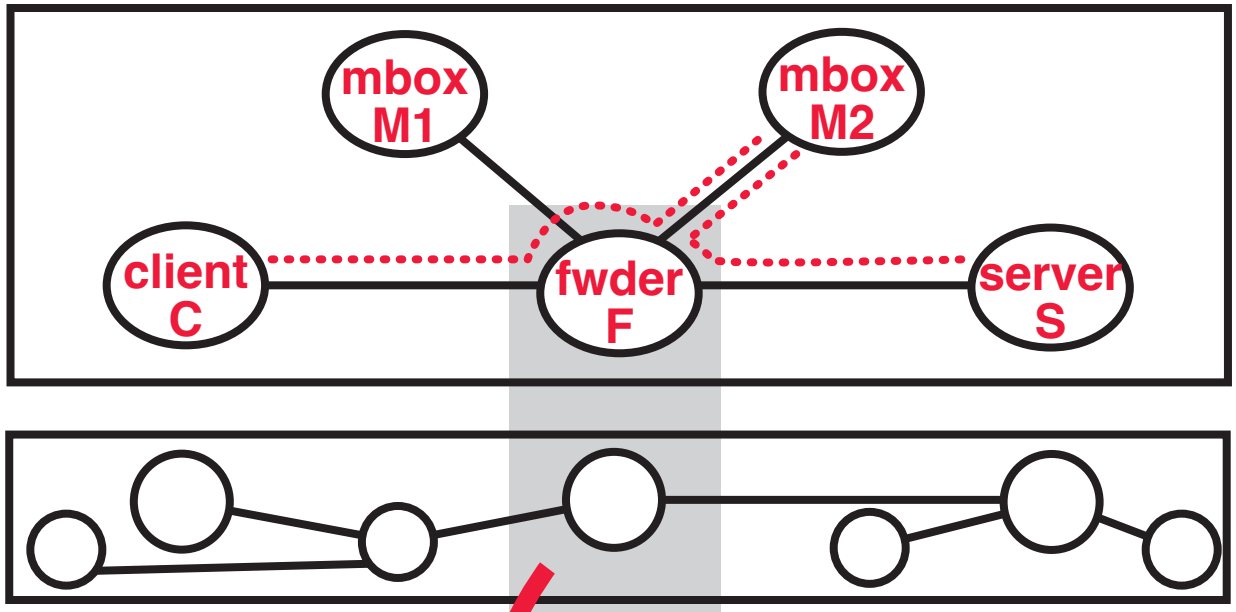


CLOUD COMPUTING 2: ENTERPRISE NETWORK

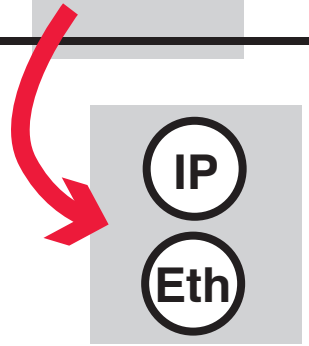
IP network . . .

. . . has naming, topology, and routing designed for enterprise . . .

. . . and is layered on an Ethernet network that implements its links



When the enterprise moves to the cloud, can the image (state) of the enterprise machine simply be loaded onto a VM?



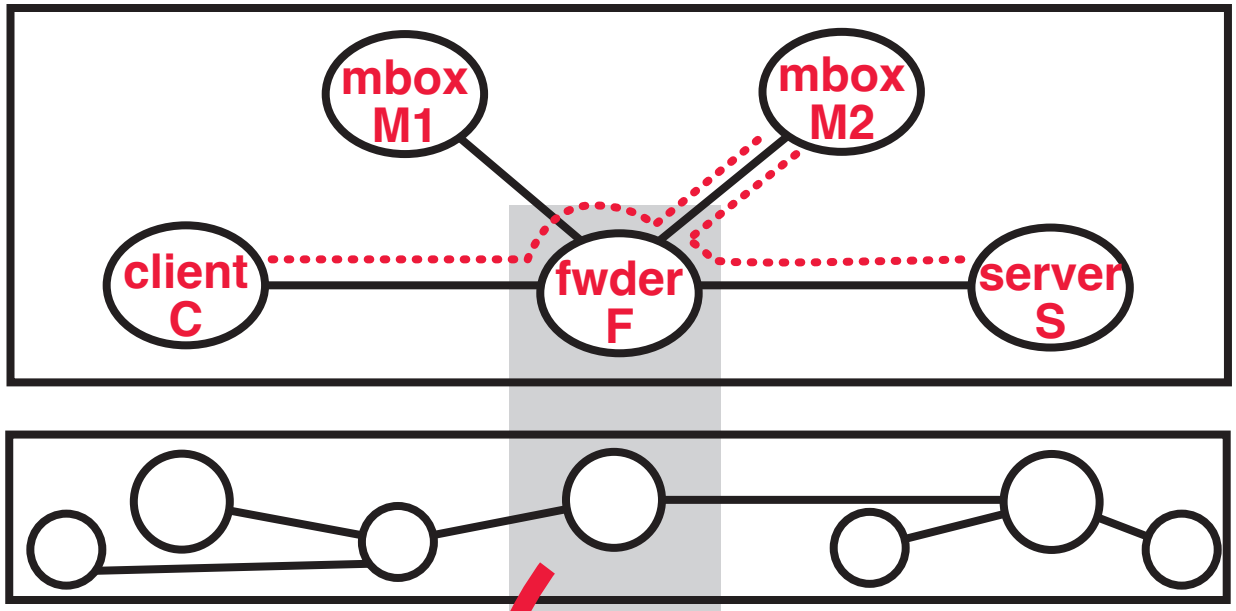
VM images are moved around the data center, for fault-tolerance and resource management

CLOUD COMPUTING 2: ENTERPRISE NETWORK

IP network . . .

. . . has naming, topology, and routing designed for enterprise . . .

. . . and is layered on an Ethernet network that implements its links



When the enterprise moves to the cloud, can the image (state) of the enterprise machine simply be loaded onto a VM?



VM images are moved around the data center, for fault-tolerance and resource management

If yes, what IP name will the VM have?

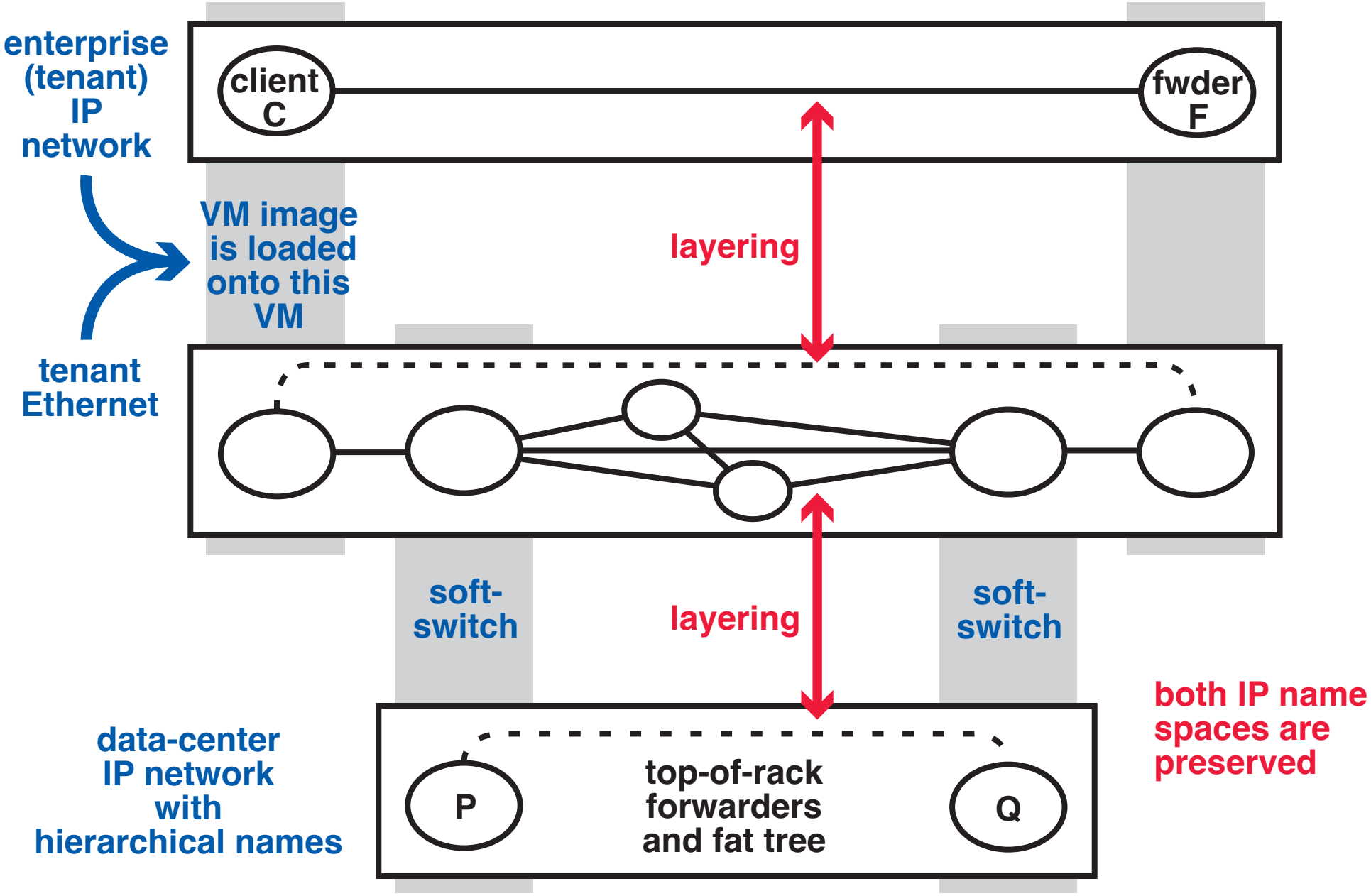
IP name in data-center network?

- now all names in enterprise network become random and change randomly
- now forwarding rules in enterprise do not work

IP name in enterprise network?

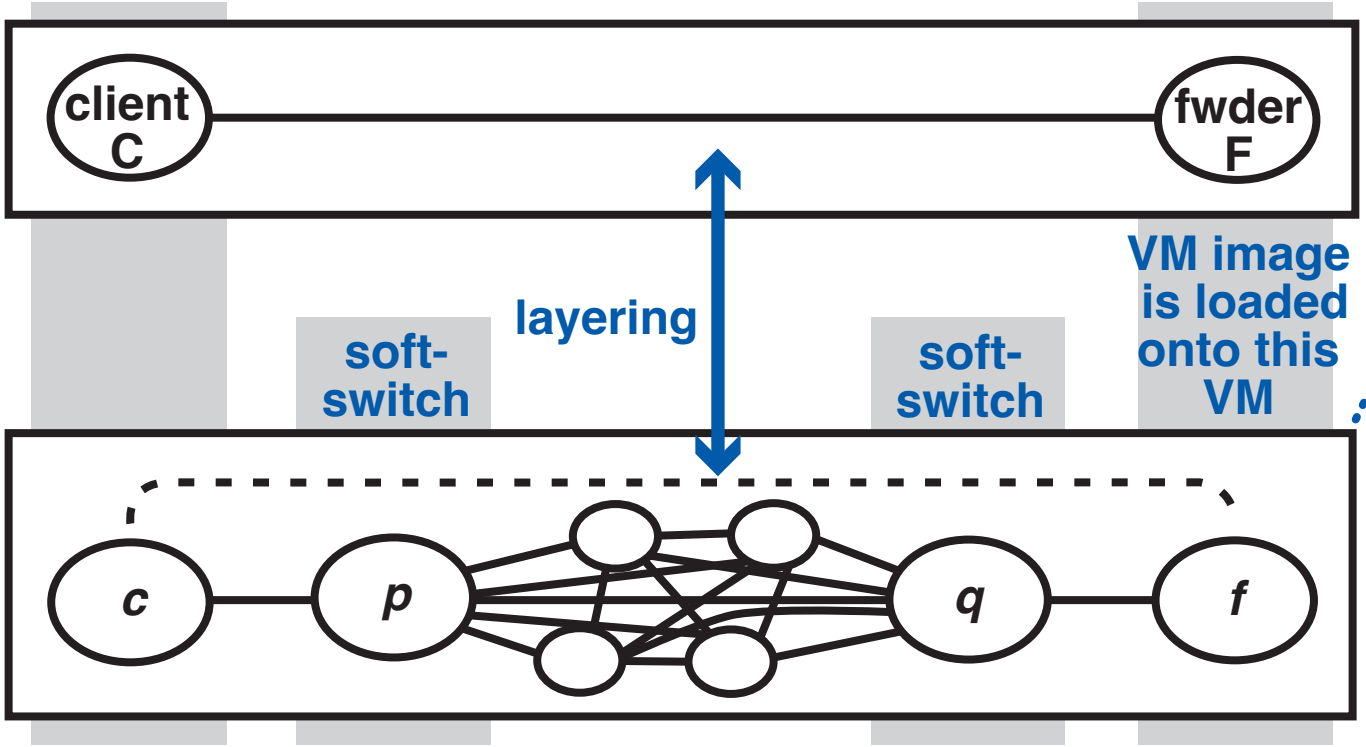
- now there are name conflicts between cloud tenants
- now names are no longer location-dependent (and change location), routing does not scale

CLOUD COMPUTING 3: A LAYERED SOLUTION



CLOUD COMPUTING 4: VM IMAGES ARE MOBILE

. . . AS VM IMAGES MIGRATE ACROSS DATA CENTER FROM VM TO VM



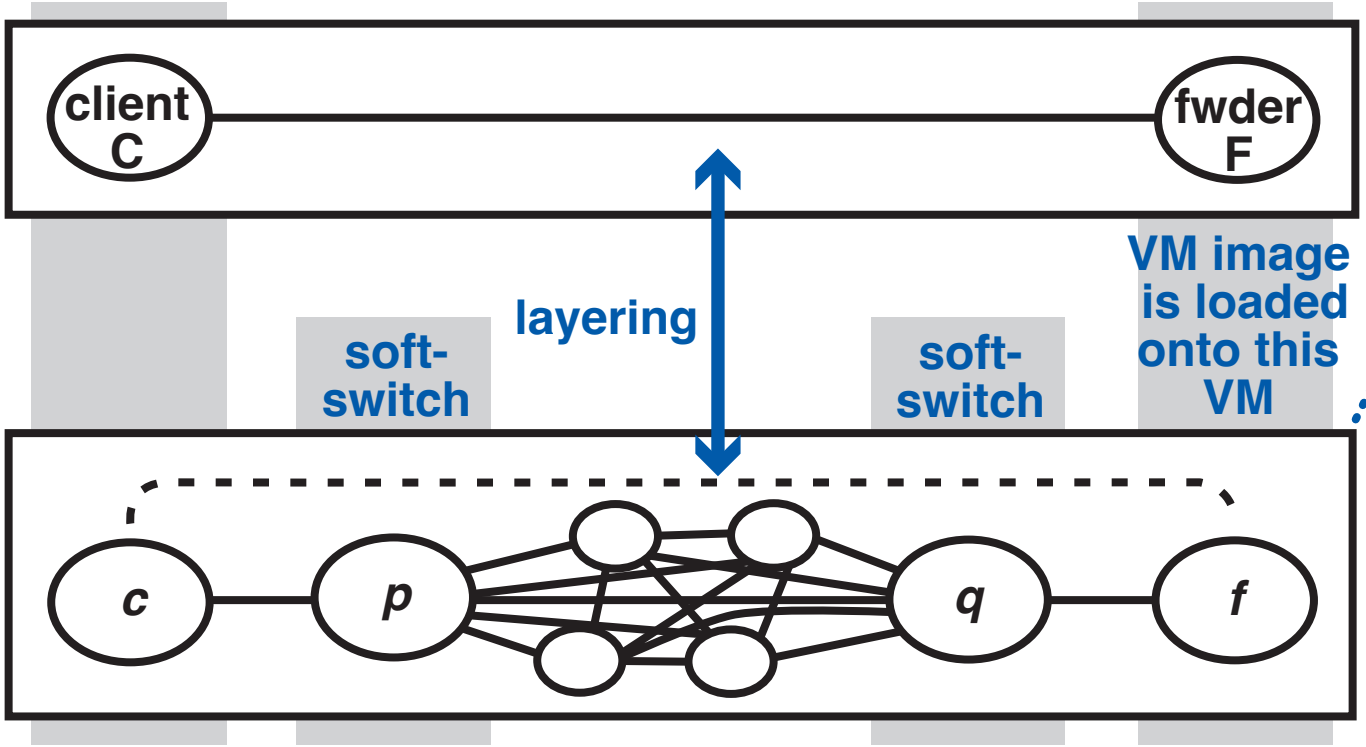
there can be many thousands of these tenant networks . . . . . with DRM in each

from *any* softswitch member such as *p*, the route to *f* is simply *q*

these routes are stored and updated in a high-throughput centralized table

CLOUD COMPUTING 4: VM IMAGES ARE MOBILE

. . . AS VM IMAGES MIGRATE ACROSS DATA CENTER FROM VM TO VM



there can be many thousands of these tenant networks . . . . . with DRM in each

from *any* softswitch member such as *p*, the route to *f* is simply *q*

these routes are stored and updated in a high-throughput centralized table

SUMMARY

evolution . . . to cloud computing
poses a problem . . . need for both routing flexibility and routing scalability
pattern solves the problem . . . with double layering

SECURITY 1: MANY ATTACKS EXPLOIT BGP

BORDER GATEWAY PROTOCOL: how the autonomous networks of the Internet exchange routing information

- critically important, complex, insecure, frequently and successfully attacked

  attackers hijack your packets, send them to the wrong place

- many people care about this, and have for a long time, but the problem remains

FOR SOME PEOPLE, THE ONLY ANSWER IS TO START OVER

SCION: a new Internet design

- does not rely on a single source of trust

  in today's world, no country or organization is trusted by everyone

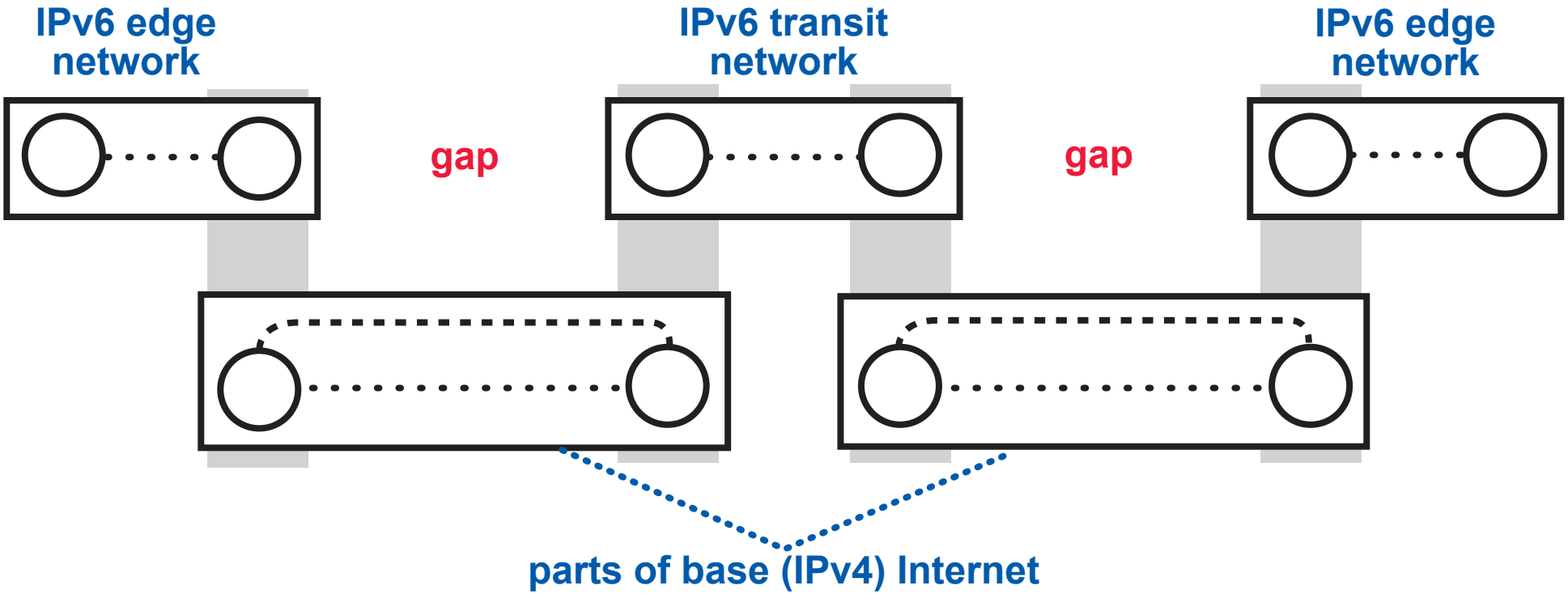
- has provably correct and secure routing (according to a stated set of properties)

- is being used in the real world, particularly by European banks

- maybe, someday, the Internet will be made of SCION networks

SECURITY 2: FOR NOW, SCION AND IP MUST CO-EXIST

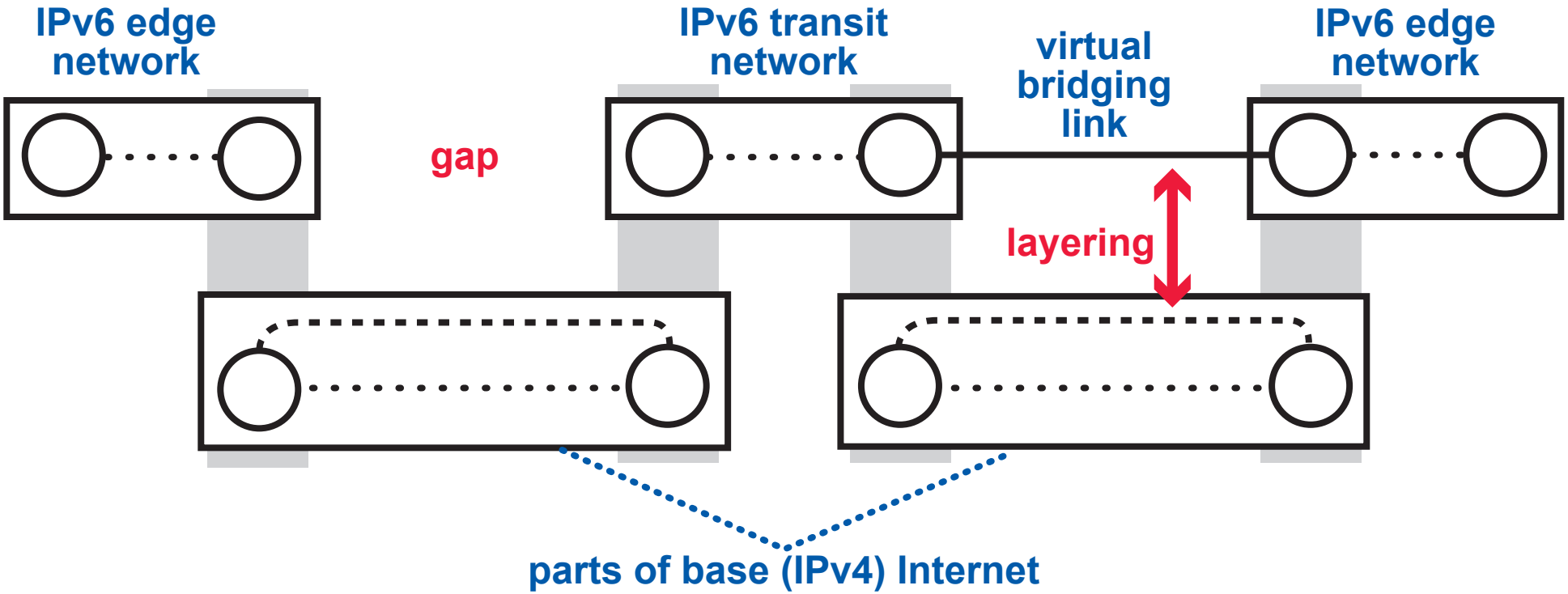
IPv4 and IPv6 already co-exist!



SECURITY 2: FOR NOW, SCION AND IP MUST CO-EXIST

IPv4 and IPv6 already co-exist!

this works where there are gaps in IPv6 coverage:

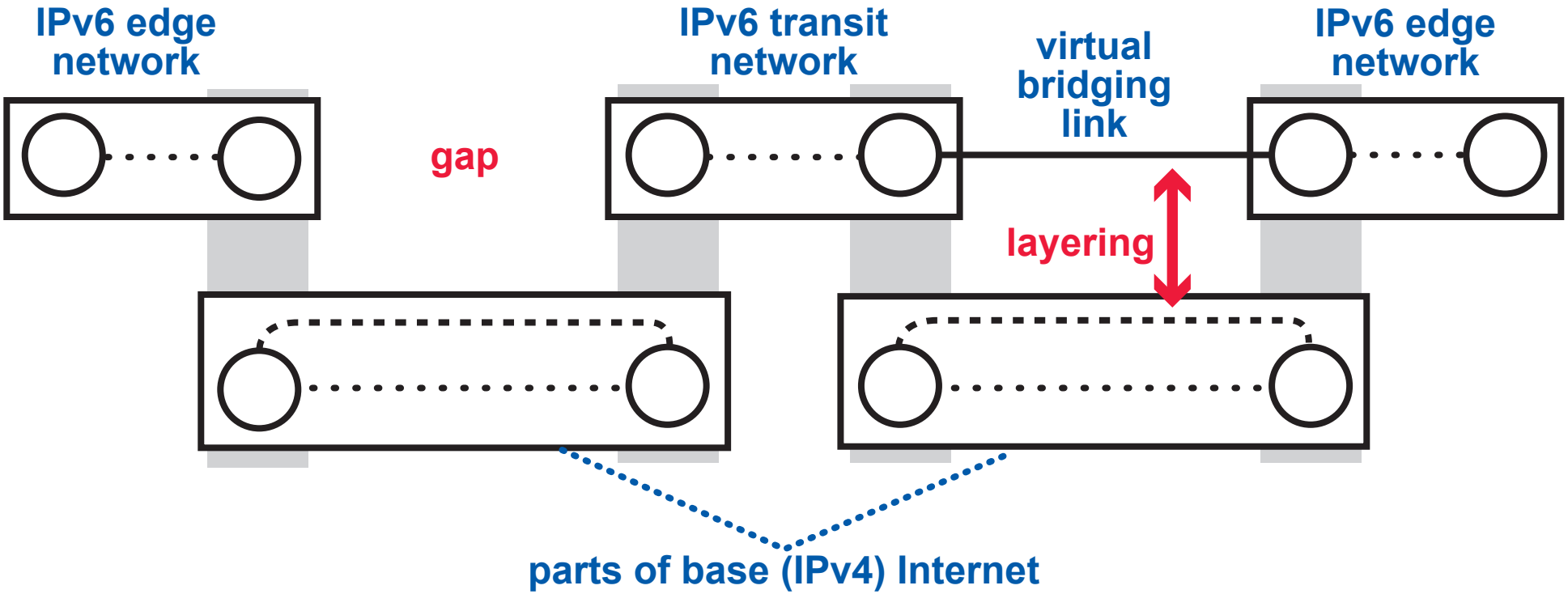


along the spans of the virtual links, IPv6 is using the resources of the IPv4 Internet, just like any other user

SECURITY 2: FOR NOW, SCION AND IP MUST CO-EXIST

IPv4 and IPv6 already co-exist!

this works where there are gaps in IPv6 coverage:



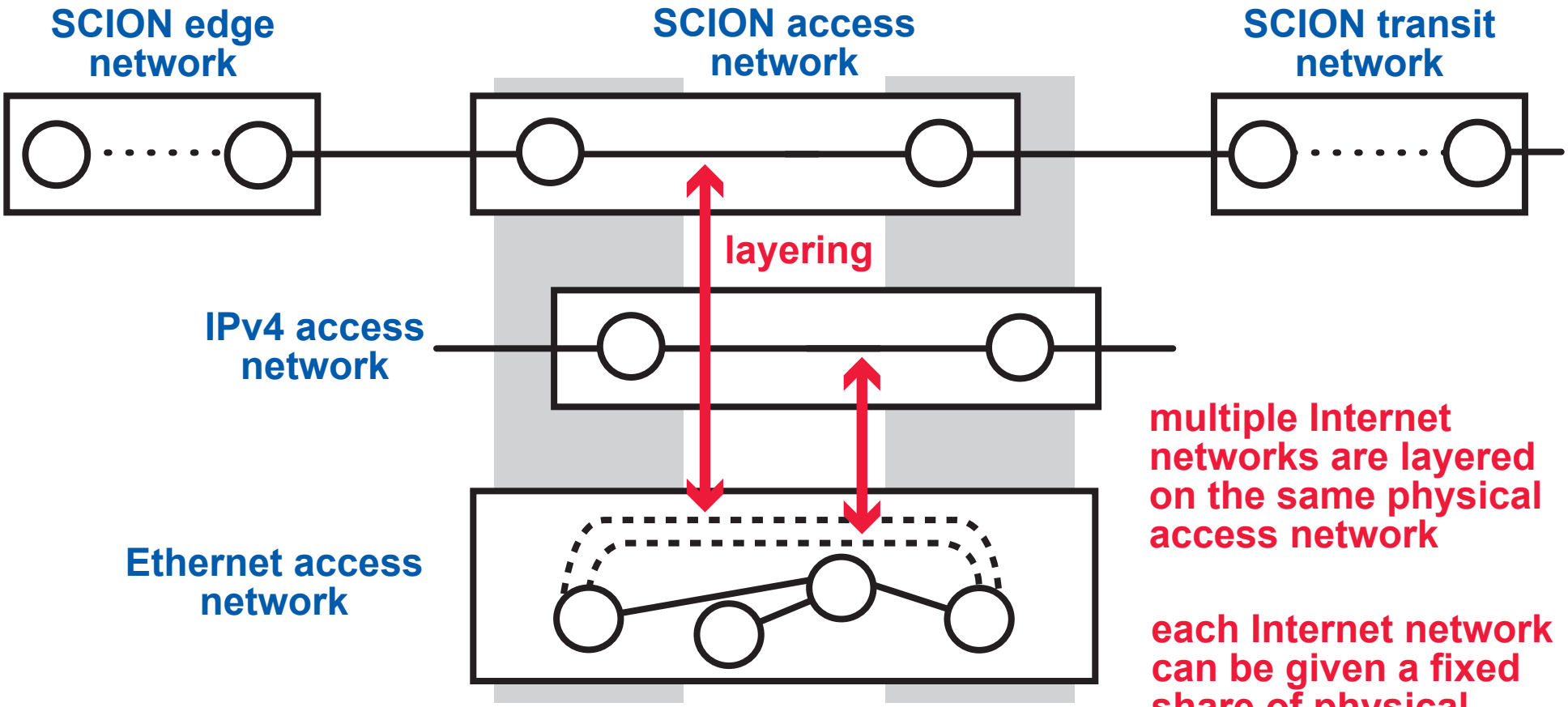
along the spans of the virtual links, IPv6 is using the resources of the IPv4 Internet, just like any other user

can SCION and IPv4 co-exist in this way?

SCION guarantees bandwidth on some network paths . . .

. . . but what if congestion or a security attack in IPv4 takes the bandwidth expected for a SCION virtual link?

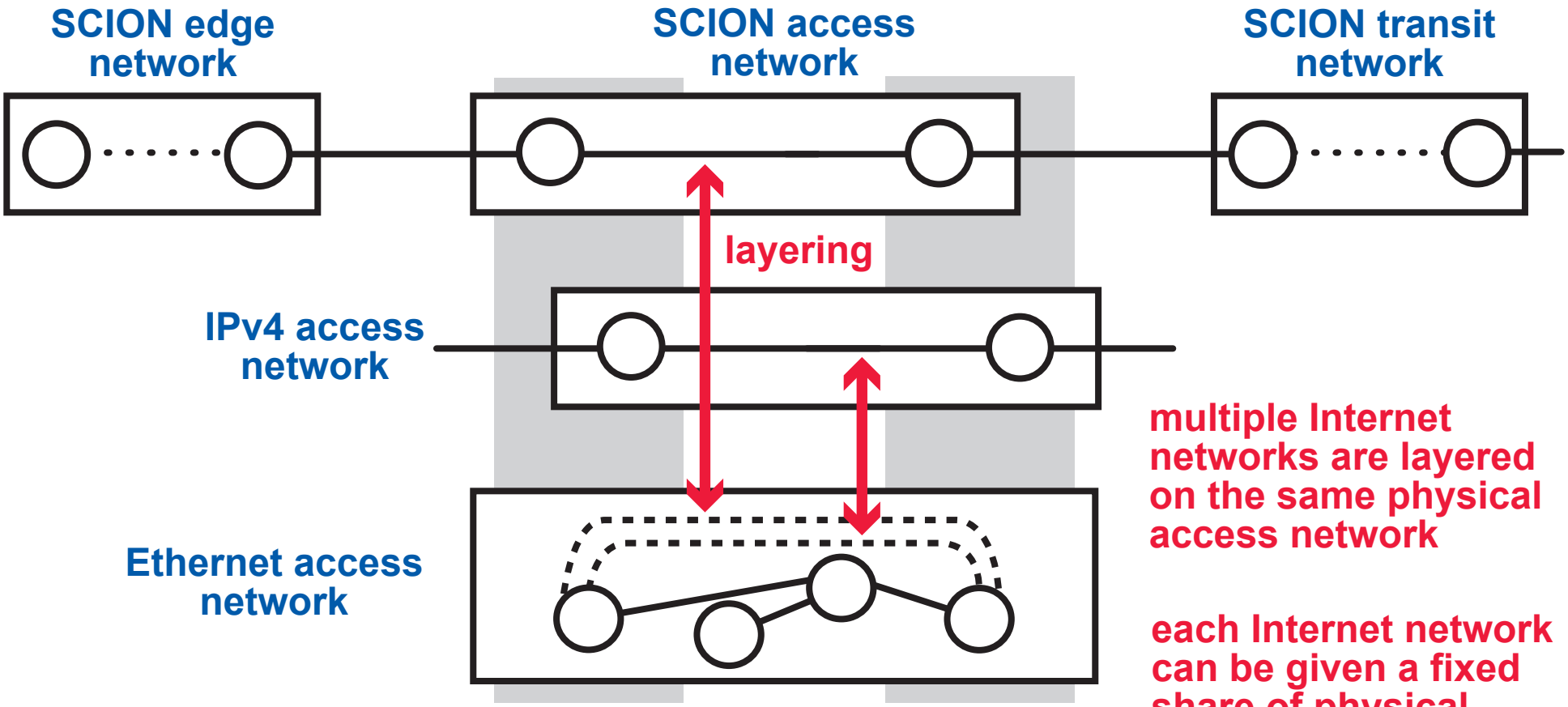
SECURITY 3: RESOURCE SLICING



multiple Internet networks are layered on the same physical access network

each Internet network can be given a fixed share of physical resources, so they are independent

SECURITY 3: RESOURCE SLICING



multiple Internet networks are layered on the same physical access network

each Internet network can be given a fixed share of physical resources, so they are independent

evolution . . . to a new Internet protocol

poses a problem . . . sharing resources, possibly in a controlled way

pattern solves the problem . . . with multiple networks layered on a single network, which may control the resources used by overlays

IN 1992, THE INTERNET WAS . . .

. . . for connecting desktop computers and mainframes

. . . for email and Web browsing

. . . mainly for nerds

. . . small and safe

. . . a sideshow to the telephone network and broadcast television

NOW, THE INTERNET IS . . .

. . . for connecting everything, but especially cellphones

. . . many people's only network

. . . the repository of most online data

. . . besieged by security threats

. . . the ubiquitous and indispensable tool of our world

IN 1992, THE INTERNET WAS . . .

. . . for connecting desktop computers and mainframes

. . . for email and Web browsing

. . . mainly for nerds

. . . small and safe

. . . a sideshow to the telephone network and broadcast television

NOW, THE INTERNET IS . . .

. . . for connecting everything, but especially cellphones

. . . many people's only network

. . . the repository of most online data

. . . besieged by security threats

. . . the ubiquitous and indispensable tool of our world

WHAT MADE THIS EVOLUTION POSSIBLE?

- progress in many technologies
- IPv4 as a general-purpose, reusable network design
- decentralized ownership and control

architecturally,

- the inherent composability of networks

  through bridging,
  layering,
  and subduction
- patterns of composition that recur

  many places in the Internet ecosystem

COMPOSITION OPERATORS IN THE REAL INTERNET ARCHITECTURE

Bridging is simple and fundamental.

Subduction exists to make Bridging and Layering more widely applicable.

The key to evolution is Layering, which has these uses in the real Internet of today:

REACHABILITY

to bridge heterogeneous networks

ROUTING

to improve the scalability and flexibility of routing

- ***improve path performance (latency, loss rate, bandwidth)***
- ***improve routing scalability***
- ***faster failure recovery***
- ***multipath routing***
- ***plug-and-play connectivity***

RESOURCES

to share or "slice" resources

- ***multi-tenant clouds***
- ***platforms for Internet experiments***
- ***protocol evolution of base Internet***

SERVICES

to provide enhanced user services

- ***mobility,***
- ***multicast,***
- ***security,***
- ***privacy,***
- ***data-centric networking***

WANT TO KNOW MORE?

THIS BOOK HAS ALL THE DETAILS

- a zoo of network designs
- hundreds of examples
- definition of enhanced network services
- more building blocks and how to use them
- research ideas

`compositionalnetarch.org`

- a recording of this talk
- the complete formal model

