THE COMPOSITIONAL ARCHITECTURE OF THE INTERNET *

Pamela Zave

Princeton University

Princeton, New Jersey, USA

* Joint work with Jennifer Rexford.

IN 1992

THE EXPLOSIVE GROWTH OF THE WORLD-WIDE WEB BEGAN

AND IN 1993

THE LAST MAJOR CHANGE WAS MADE TO THE "CLASSIC" INTERNET ARCHITECTURE

THE LAST MAJOR CHANGE: CLASSLESS ADDRESSING

BEFORE

a router can advertise a network of size 256 (Class C)

with a 24-bit address

a router can advertise a network of size 65,636 (Class B)

with a 12-bit address

a router can advertise a network of size 16,777,216 (Class A)

with an 8-bit address

AFTER

a router can advertise a network of size $2^{(32 - X)}$

with an X-bit address “. . . /X”

WHAT HAS HAPPENED SINCE 1993?

- most of the world's . . .

  . . . telecommunication infrastructure

  . . . entertainment distribution . . .

  has moved to the Internet

- an explosion of security threats
- most networked devices are mobile
- cloud computing
- exhaustion of the IP address space
- the need for elastic resource allocation instead of over-provisioning

A CONUNDRUM:

The “classic” Internet architecture (how experts describe the Internet) has not changed since 1993, . . .

. . . yet the Internet has met all these new challenges, at least to some extent.

THE "CLASSIC" INTERNET ARCHITECTURE

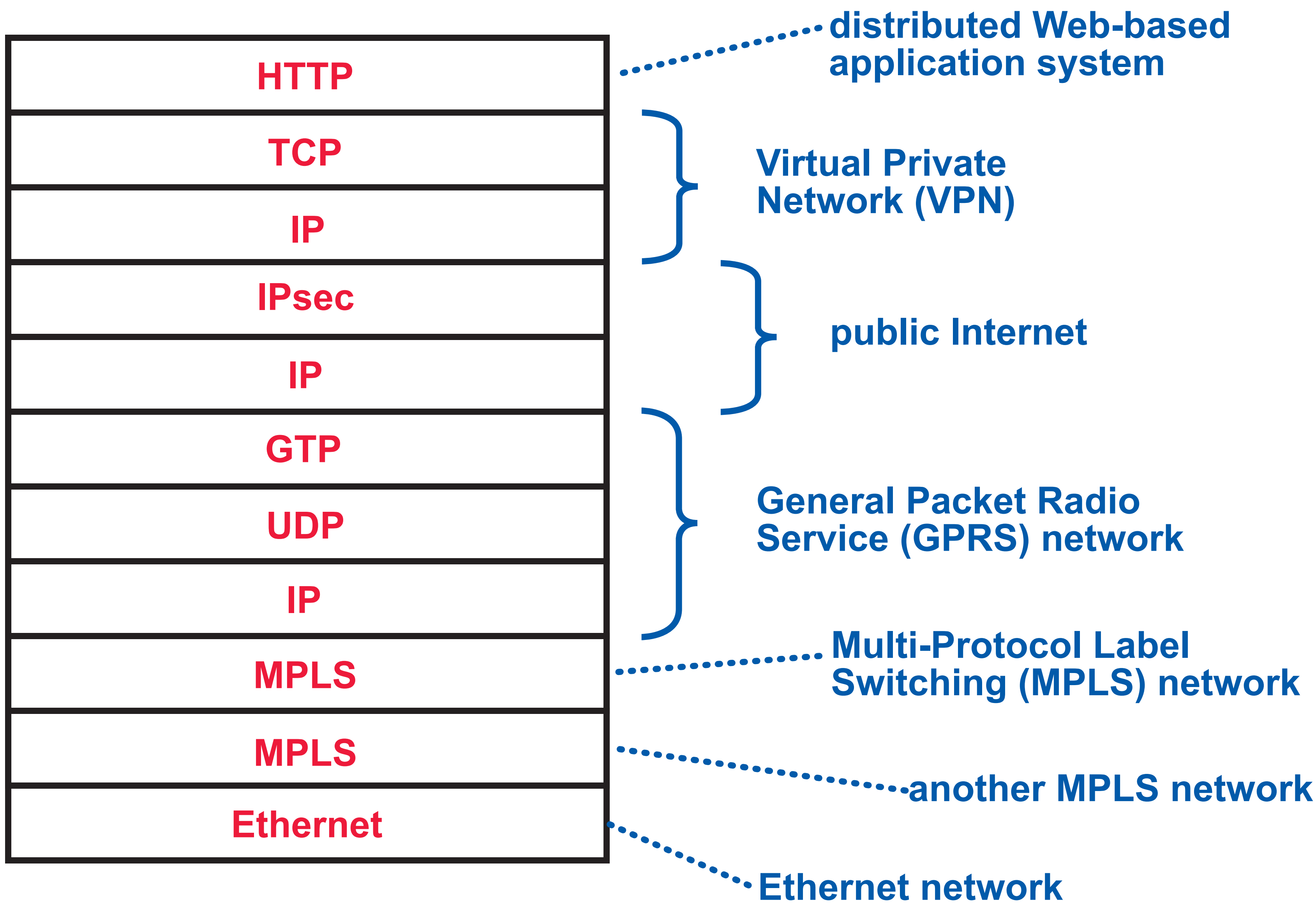
APPLICATION LAYER	applications and mnemonic names
TRANSPORT LAYER	reliable byte streams, datagrams
NETWORK LAYER	best-effort global packet delivery
LINK LAYER	best-effort local packet delivery
PHYSICAL LAYER	diverse physical media (wires, optical fibers, radio channels)

so we expect
a typical packet
to look like this

HTTP header
TCP header
IP header
Ethernet header

THE REALITY: THIS IS A TYPICAL PACKET IN THE AT&T BACKBONE

packets sampled elsewhere
would look different, but
might be equally complex



WHY WE NEED A BETTER MODEL . . .

. . . instead of just talking about the classic Internet architecture

and saying "there are a lot of exceptions"

IT WOULD BE NICE TO KNOW . . .

- How *has* the Internet evolved to meet the new challenges?
- How *should* it evolve in the future?

so far, efforts to design "future Internet architectures" have convinced no one

PROGRAMMABILITY

- After 25 years of hard work by the networking community, networks are now programmable.
- But there has been much less progress in knowing what to program.

as we all know, you can make a bigger mess with software than you can with hardware

SECURITY

- Security attacks are unforgiving—details and exceptions cannot be ignored.
- Verification of trustworthy network services requires a more holistic approach.

A BETTER MODEL: THE INTERNET IS A FLEXIBLE COMPOSITION OF MANY NETWORKS

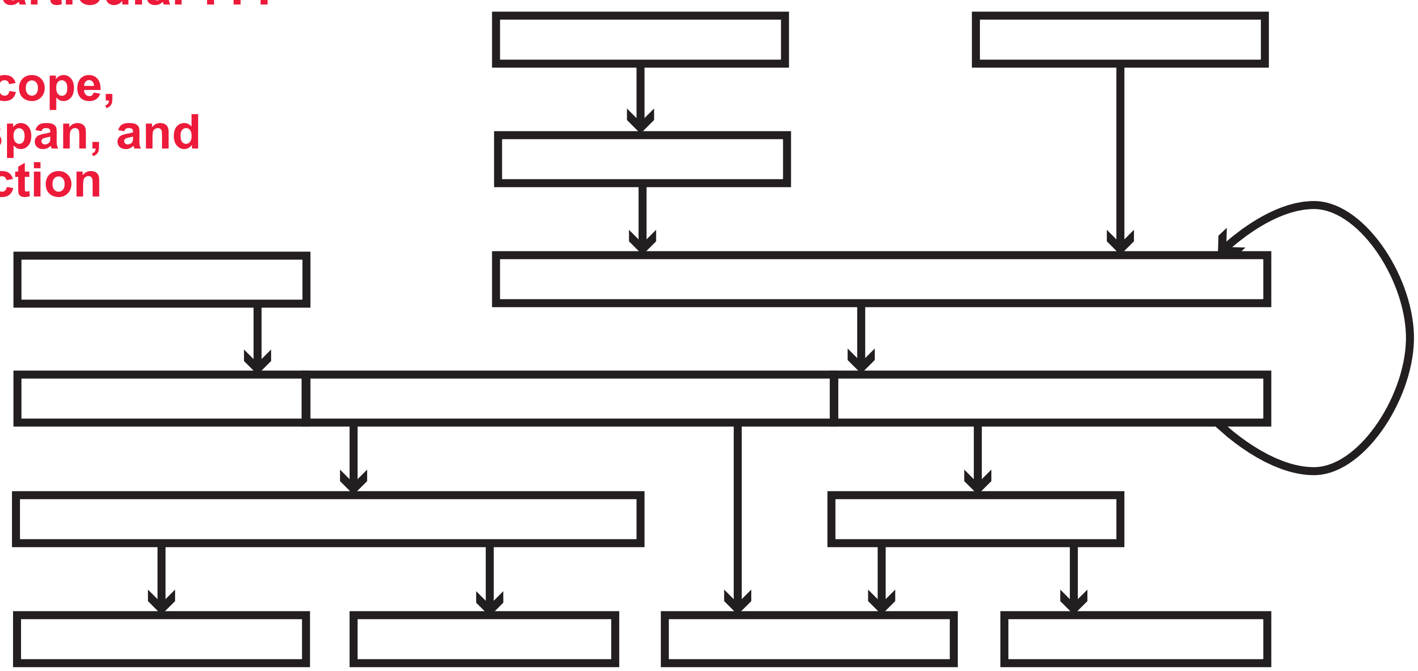
global networking
as we know it

many more than those acknowledged
in the classic architecture

each network has all the same
basic mechanisms, . . .

. . . but in each network they are
specialized for a particular . . .
. . . purpose,
. . . membership scope,
. . . geographical span, and
. . . level of abstraction

because all networks have
fundamental similarity, they all have
common interfaces for composition



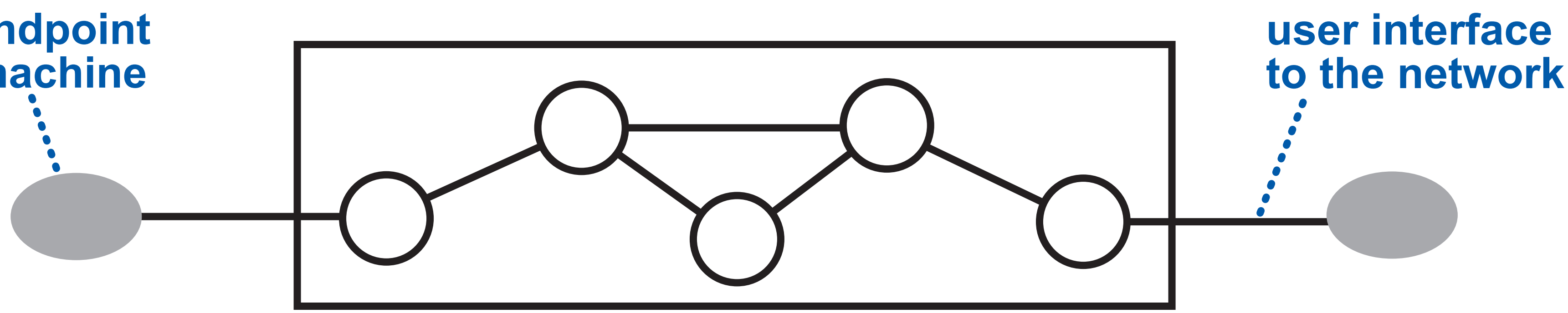
the Internet protocol suite implements a general-purpose network design and
is available on most networked devices—so it is re-used for many purposes

OLD: THE END-TO-END PRINCIPLE

The functions of a network should be minimized, so that it serves everyone efficiently, . . .

. . . and whenever possible, services should be implemented in endpoint machines.

or, "smart edge, dumb network"



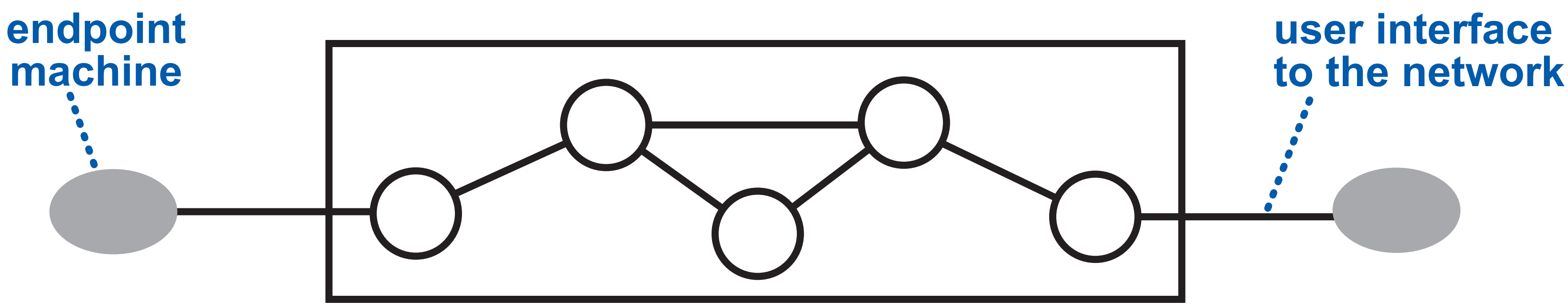
the End-to-End Principle is a design principle, but it has been so influential that it is assumed to be descriptive

OLD: THE END-TO-END PRINCIPLE

The functions of a network should be minimized, so that it serves everyone efficiently, . . .

. . . and whenever possible, services should be implemented in endpoint machines.

or, "smart edge, dumb network"



the End-to-End Principle is a design principle, but it has been so influential that it is assumed to be descriptive

today there are many exceptions:

- many services are implemented inside the network, . . .

  . . . by middleboxes and programmable routers

- cannot control network congestion without the cooperation of endpoints

today we know . . .

. . . that if we want to verify network services . . .

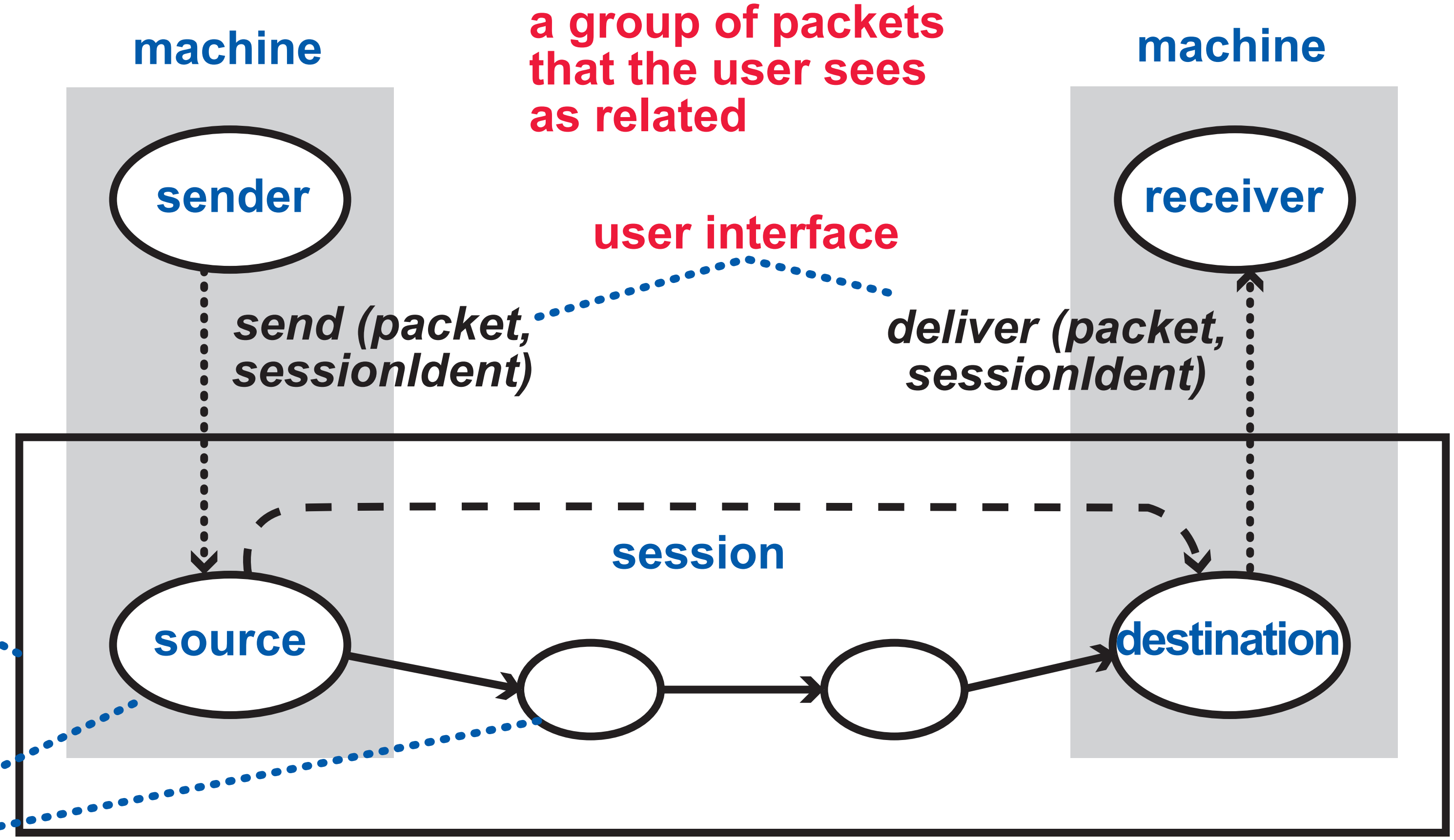
. . . we must include in our model all the agents involved in providing those services

NEW: USER INTERFACES ARE INSIDE MACHINES

the user of a network is a distributed application system— its modules must communicate through network services

an instance of network service is a session; a session transmits a group of packets that the user sees as related

modules on the same machine communicate through its operating system or hardware



a member of a network is a software or hardware module that implements some of the network protocols

OLD: LAYERS ARE FIXED, HAVE DISTINCT FUNCTIONS

classic Internet architecture has 5 layers, OSI model has the same 5 plus 2 others

routing is the control mechanism that chooses packet paths and encodes paths in forwarding tables

forwarding is the mechanism that pushes packets along their paths

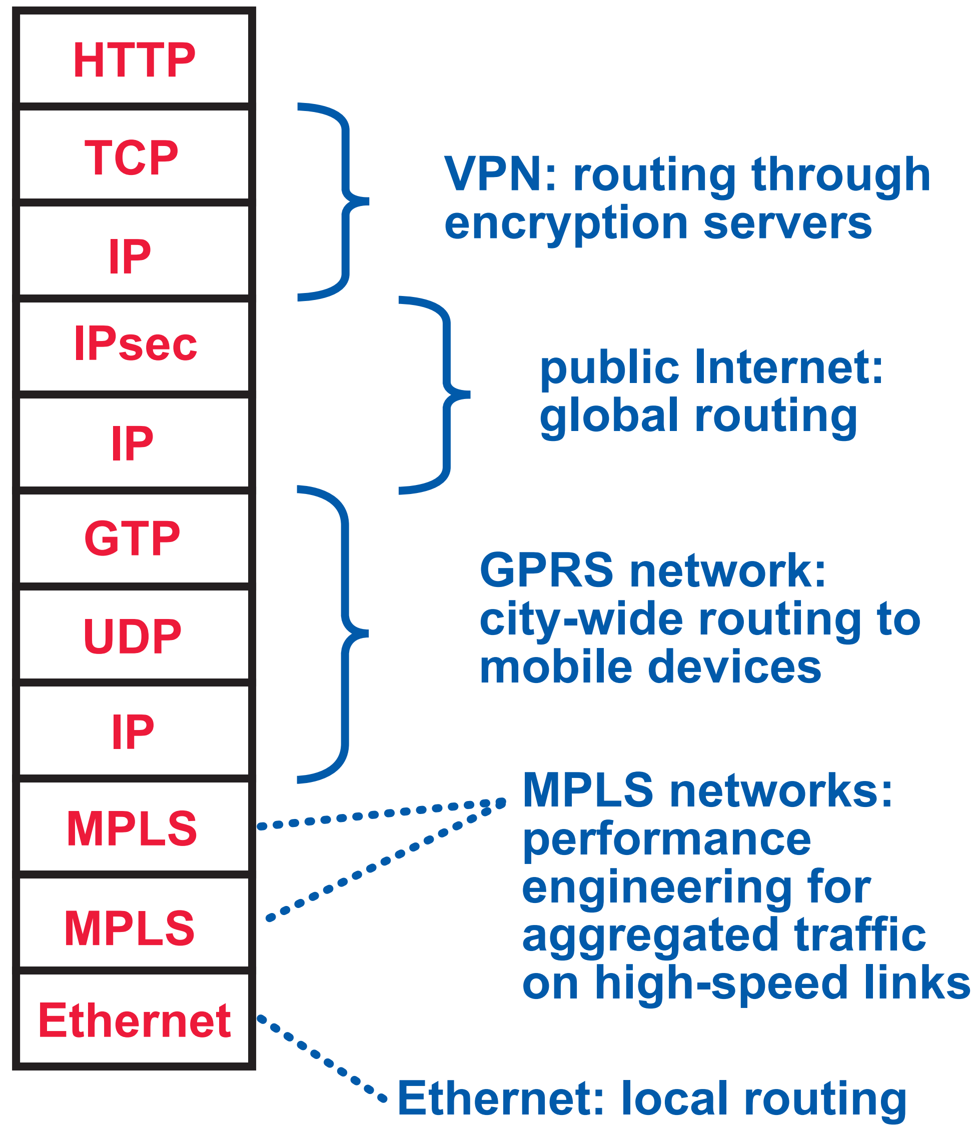
in both reference architectures, there is routing and forwarding only in the link layer (local) and network layer (global)

in this realistic example, there is routing and forwarding in each of the six networks, . . .

. . . with different purposes,

. . . over different spans,

. . . allocating different resources

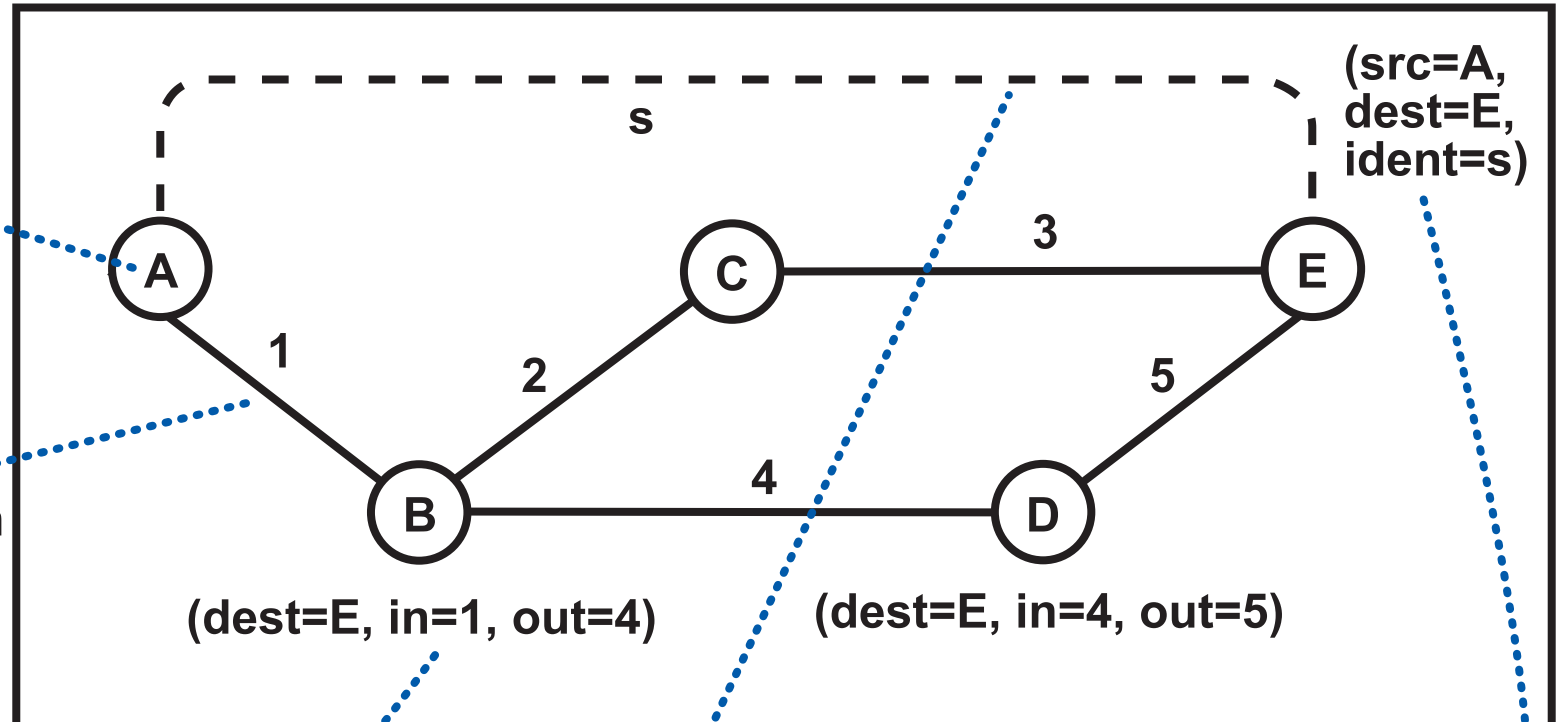


NEW: LAYERS IN A COMPOSITION HIERARCHY ARE SELF-CONTAINED NETWORKS

each network is a microcosm of networking
with all of the basic mechanisms, . . .
. . . all of which can be specialized,
. . . and some of which can be vestigial

members have names from a **namespace**

members are connected by **links** (communication channels)



routing chooses packet paths and populates **forwarding tables**, which are used by the **forwarding protocol**

a **session** is an instance of network service

the service is implemented by a **session protocol**, with **session state** in members

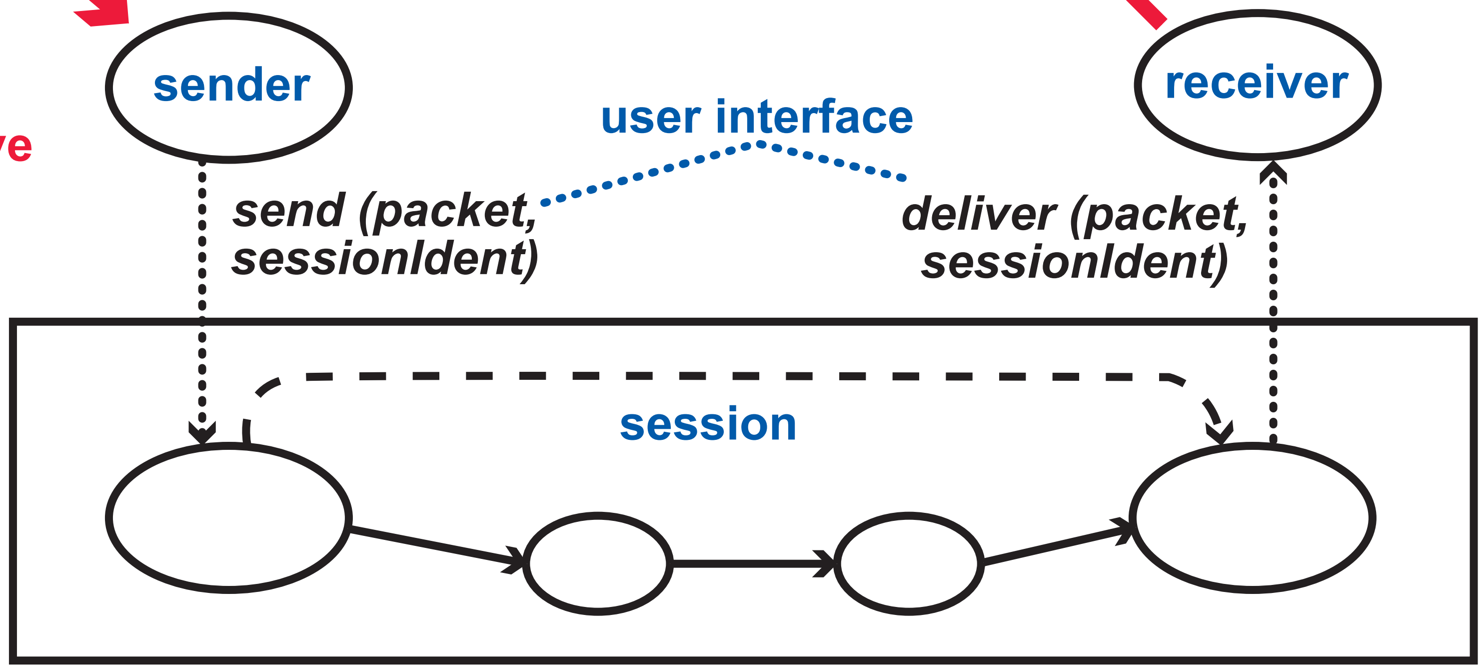
REQUIREMENTS ON NETWORKS

The users put a load (of sessions and packets) on the network.

The network delivers communication services with desirable properties.

A network has a single administrative authority . . .

. . . that is responsible for its services.



REACHABILITY

- what are the possible destinations?

PERFORMANCE

- maximum latency
- minimum bandwidth
- packet loss rate
- availability

SERVICE-SPECIFIC BEHAVIOR

- interoperation
- synchronization
- guaranteed, ordered delivery
- load-balancing
- session persistence despite endpoint mobility

SECURITY

- access control
- DoS protection
- authentication
- privacy
- data integrity
- law enforcement

SELF-CONTAINED REASONING ABOUT A NETWORK

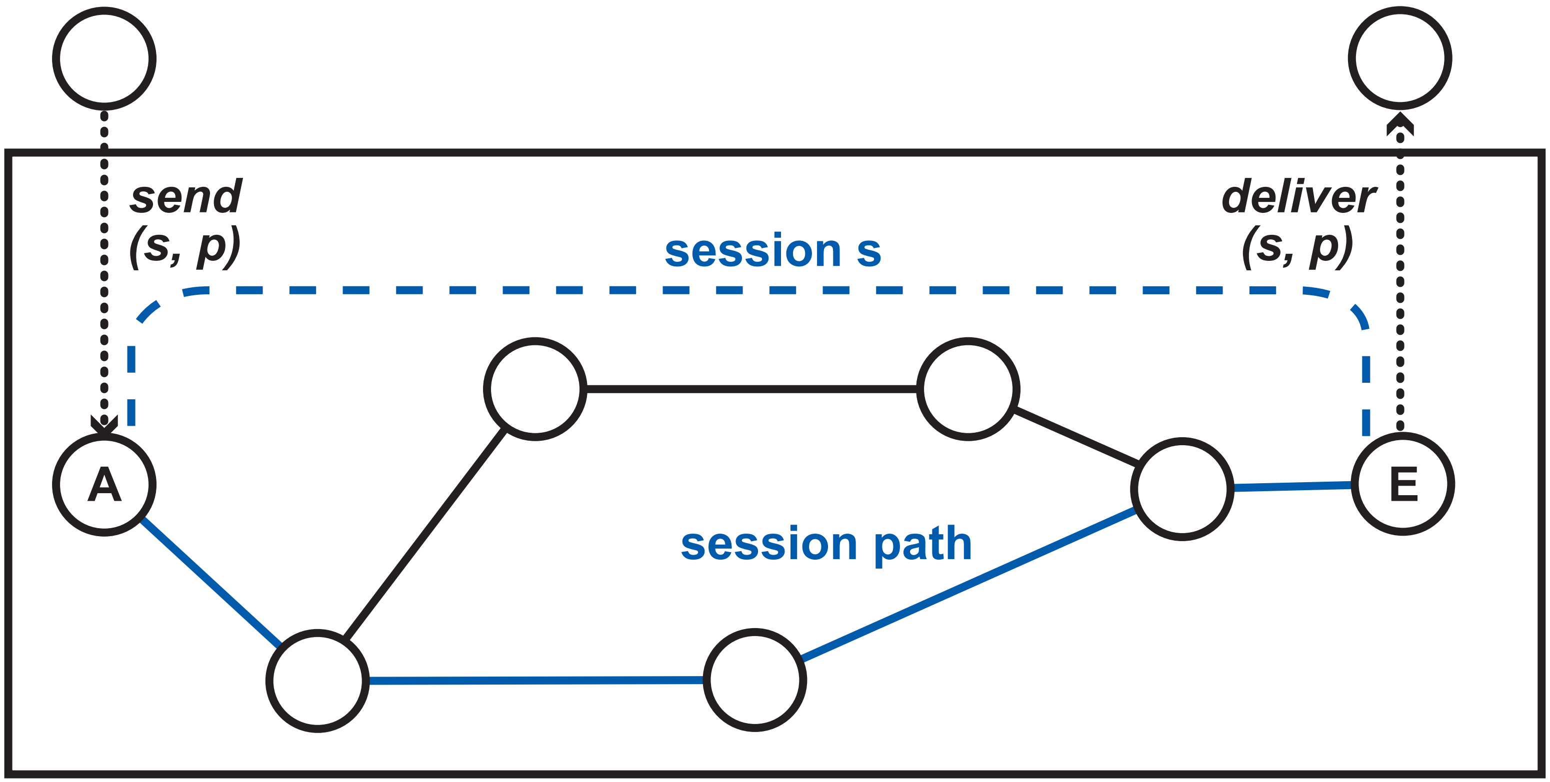
REACHABILITY

which members can be reached from A?

SESSION PERFORMANCE

what is the minimum bandwidth, maximum latency?

PROTOCOLS



SECURITY

is E protected from DoS attacks and malware?

SELF-CONTAINED REASONING ABOUT A NETWORK

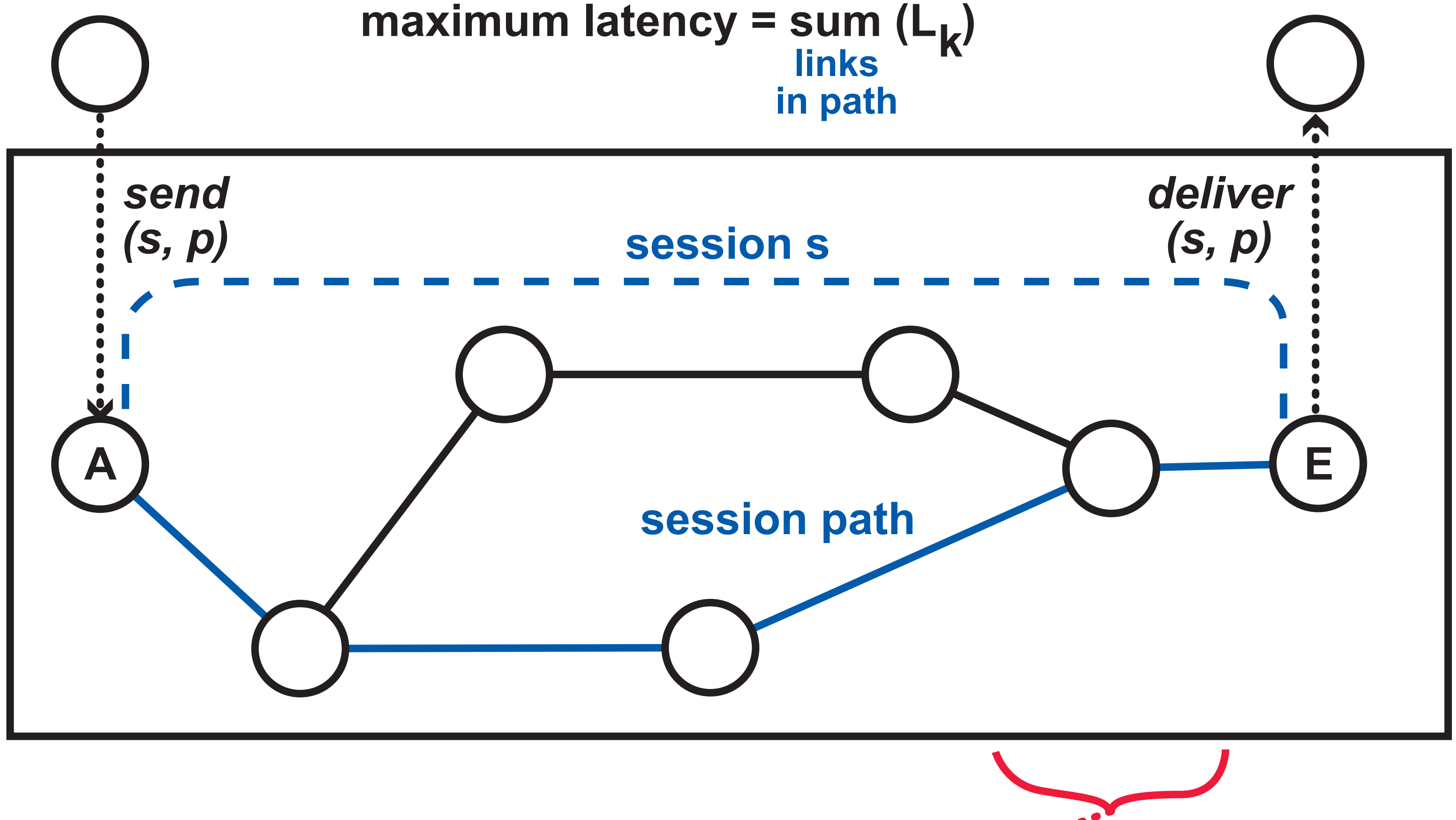
REACHABILITY

reachability from A is the transitive closure of the forwarding relation

SESSION PERFORMANCE

minimum bandwidth = min $(S_k(B_k))$
links in path
session's share of bandwidth

maximum latency = sum (L_k)
links in path



PROTOCOLS

reasoning about control and session protocols

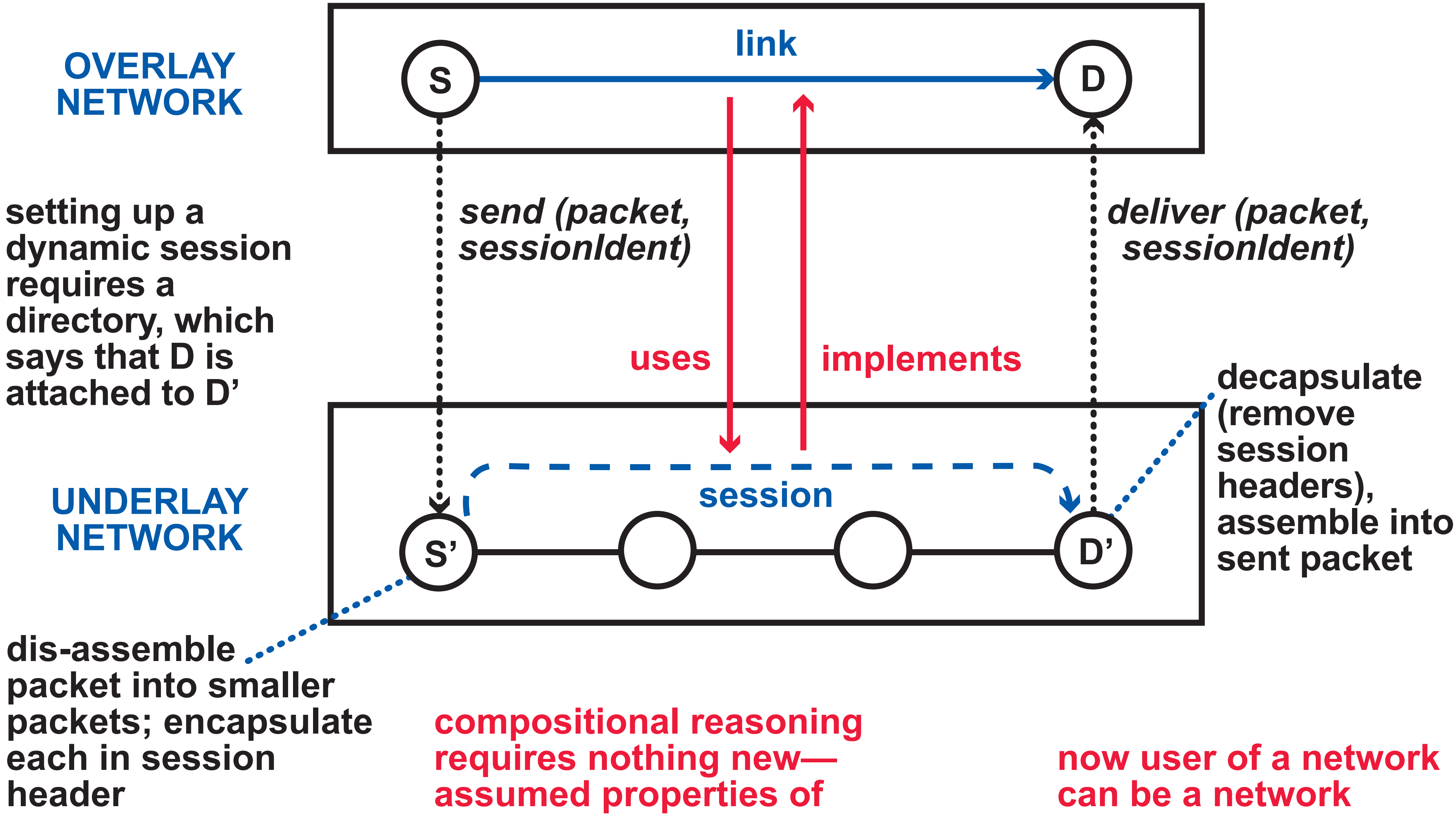
reasoning often requires assumptions about the behavior of links

SECURITY

all paths to E go through middleboxes that protect it from DoS attacks and malware

A COMPOSITION OPERATOR: LAYERING

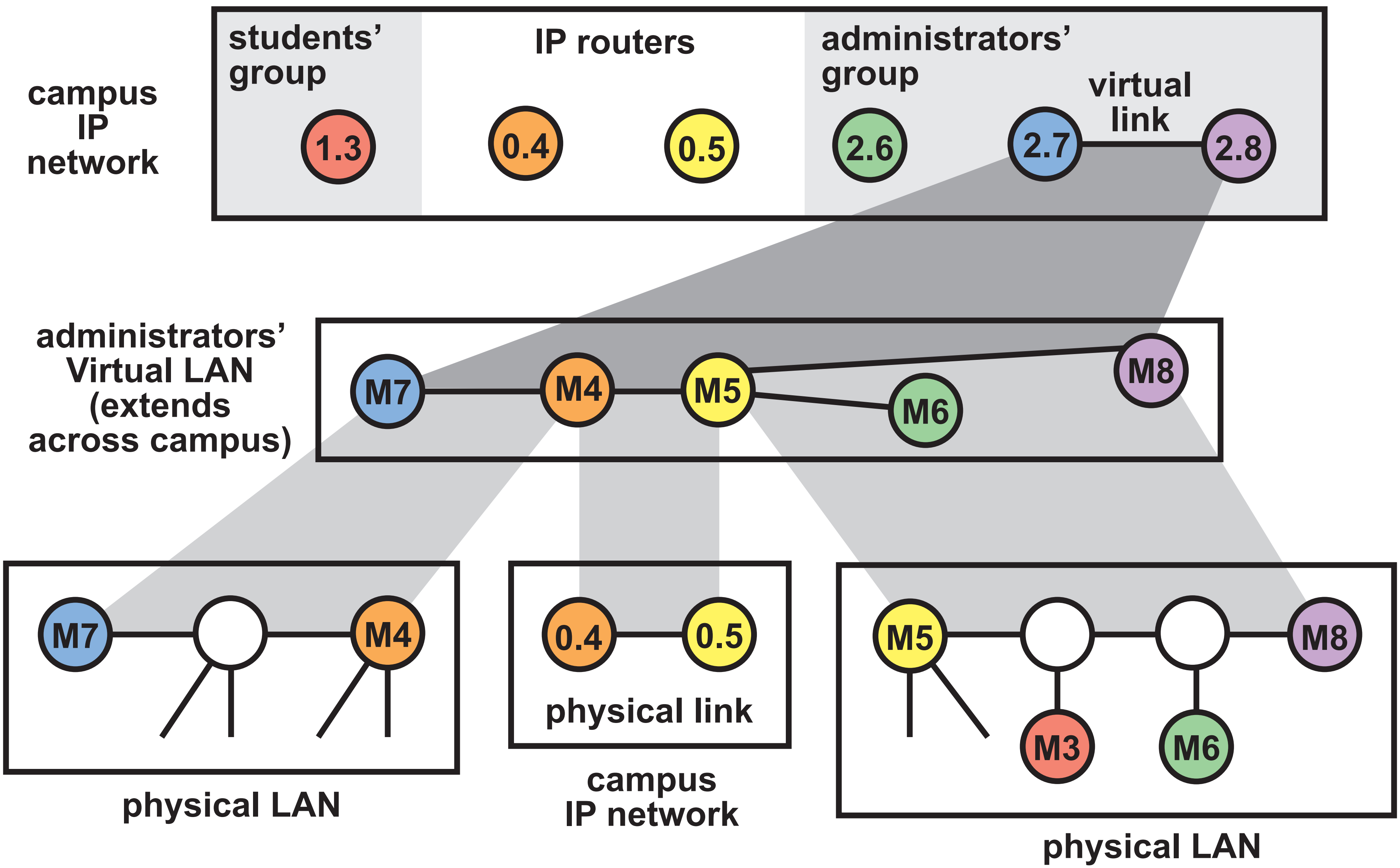
a link in an "overlay" network . . . is implemented by a session in an "underlay" network

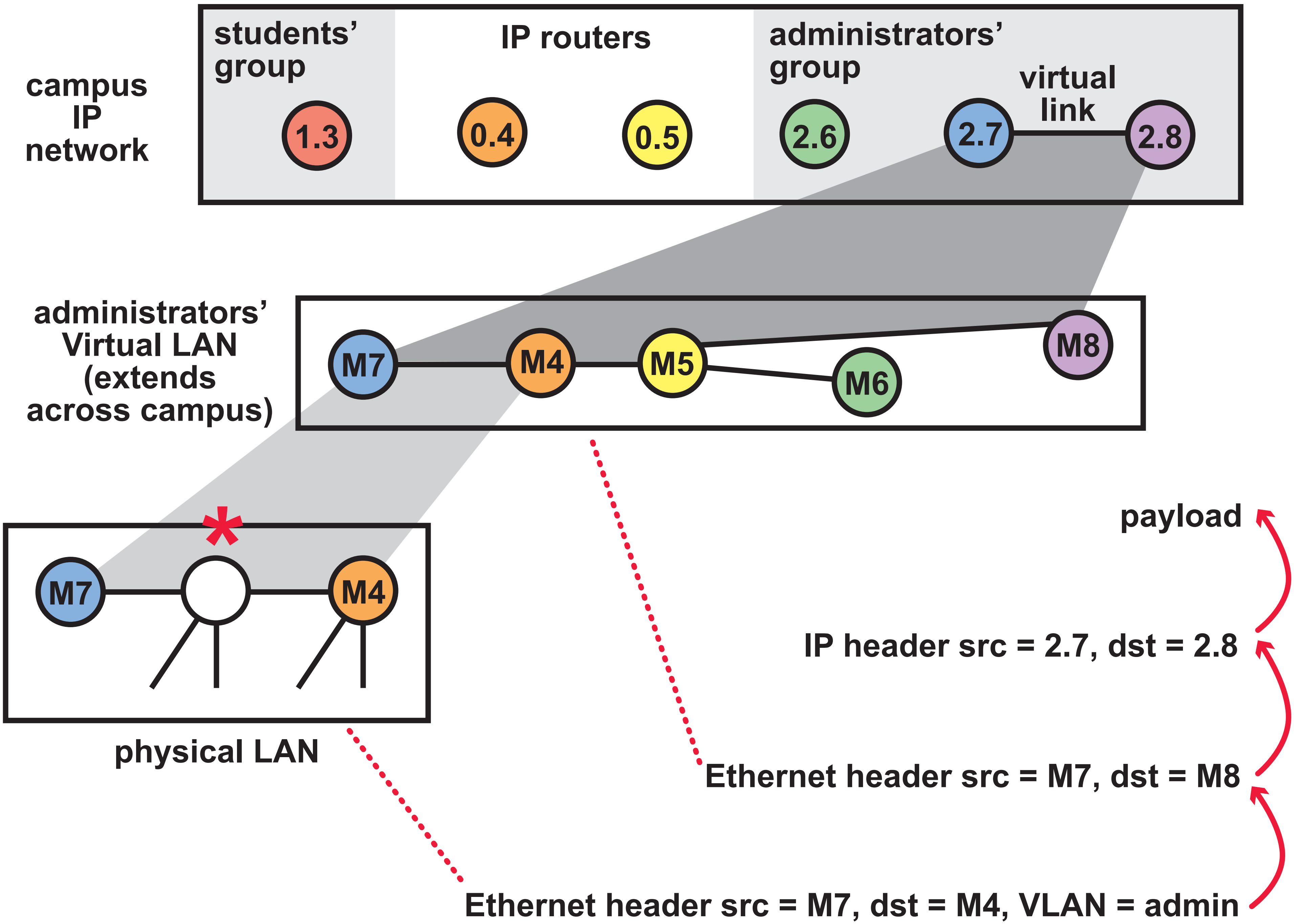


compositional reasoning requires nothing new— assumed properties of overlay link are specified properties of underlay session

now user of a network can be a network instead of a distributed application system

LAYERING NOT IN CLASSIC ARCHITECTURE:
CAMPUS NETWORK WITH VLANS FOR SECURITY



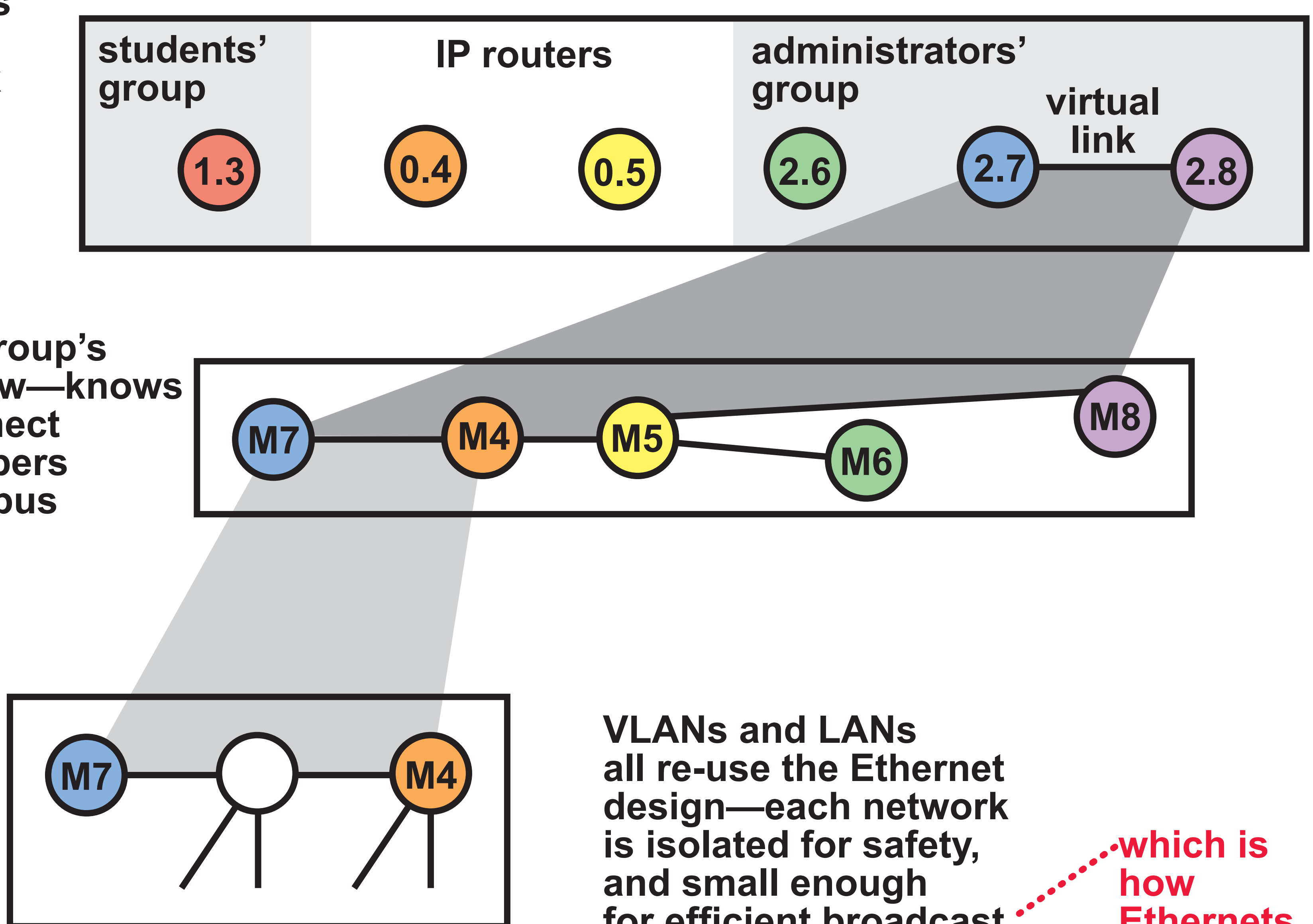

campus IP network
students' group
1.3
IP routers
0.4
0.5
administrators' group
2.6
2.7
virtual link
2.8
administrators' Virtual LAN (extends across campus)
M7
M4
M5
M6
M8
*
M7
M4
physical LAN
payload
IP header src = 2.7, dst = 2.8
Ethernet header src = M7, dst = M8
Ethernet header src = M7, dst = M4, VLAN = admin

WHY? TWO VIEWS OF SAME NETWORK, WITH DIFFERENT TOPOLOGIES

IP network is the security view—prefix identifies a machine's group

VLAN is a group's physical view—knows how to connect group members across campus

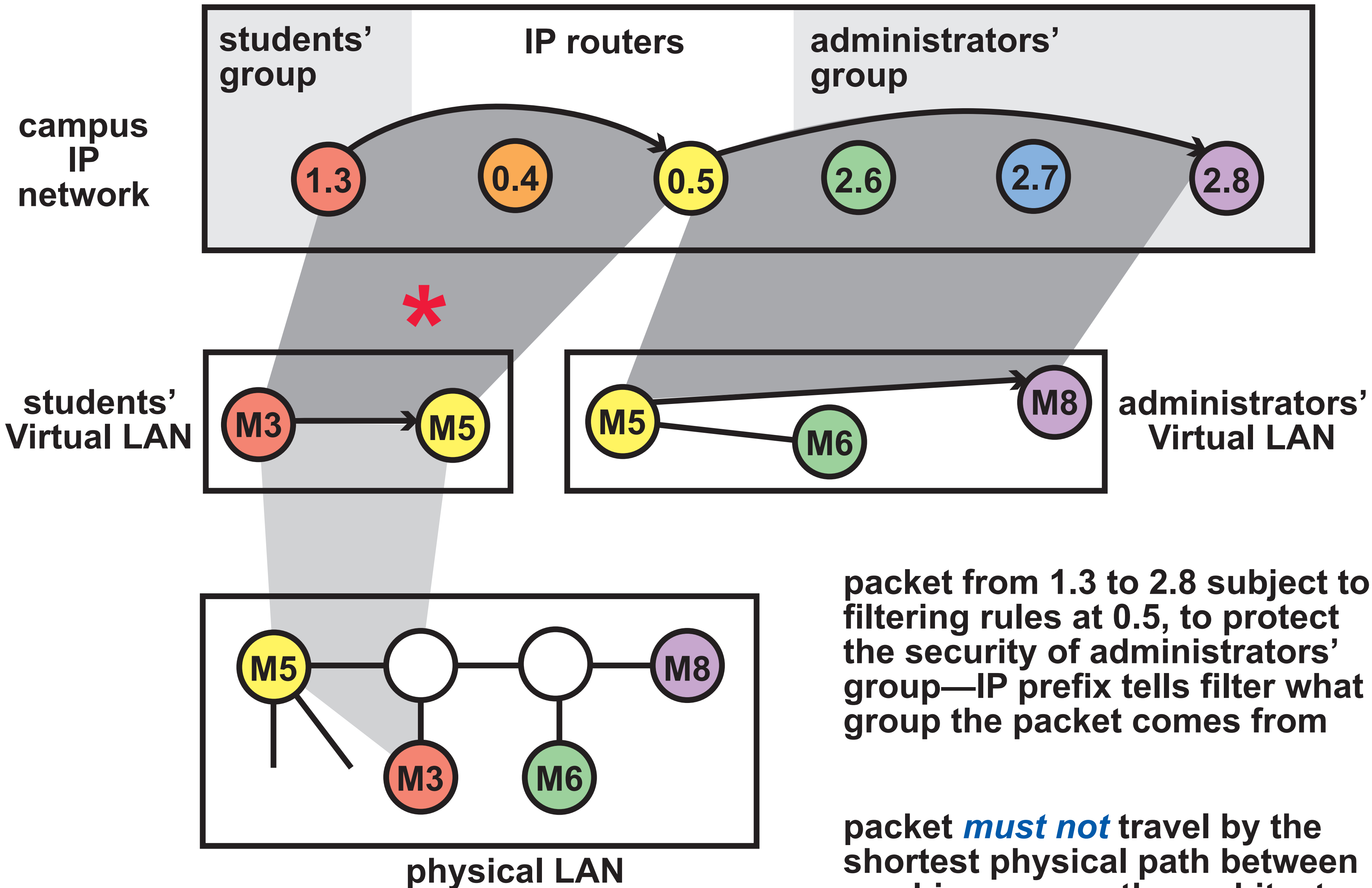
each LAN is local



VLANs and LANs all re-use the Ethernet design—each network is isolated for safety, and small enough for efficient broadcast

which is how Ethernets do routing

VERIFICATION OF INTER-GROUP SECURITY



packet from 1.3 to 2.8 subject to filtering rules at 0.5, to protect the security of administrators' group—IP prefix tells filter what group the packet comes from

packet ***must not*** travel by the shortest physical path between machines . . . so the architecture ***must*** be implemented correctly!

ANOTHER COMPOSITION OPERATOR: BRIDGING

bridging allows services to be implemented by networks chained end-to-end

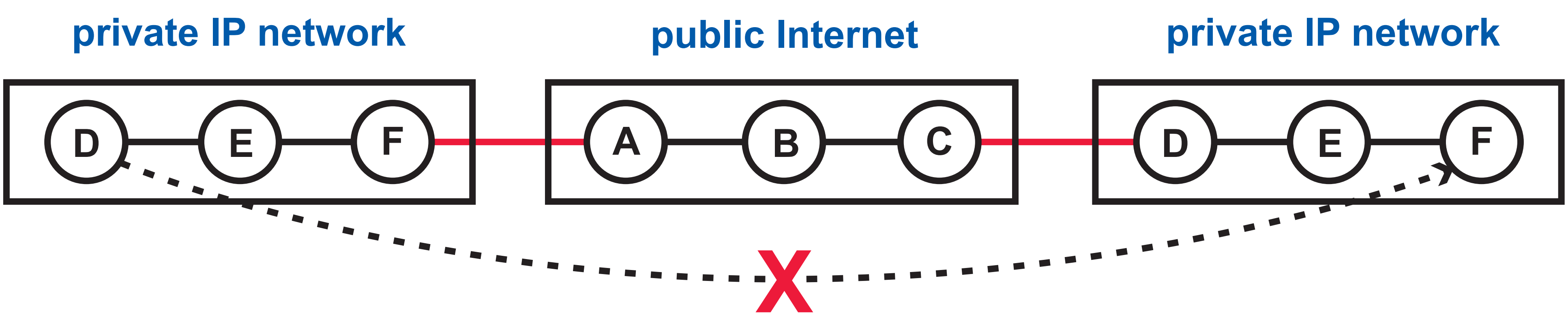
THE EASY WAY

networks have . . .
. . . same namespace
. . . same protocols
. . . globally unique names
. . . access to other networks' routing and directories

this is how the networks of the public Internet are composed—they differ only in their administrative authorities

THE HARD WAY

some constraints above do not hold, e.g., private IP networks re-use names

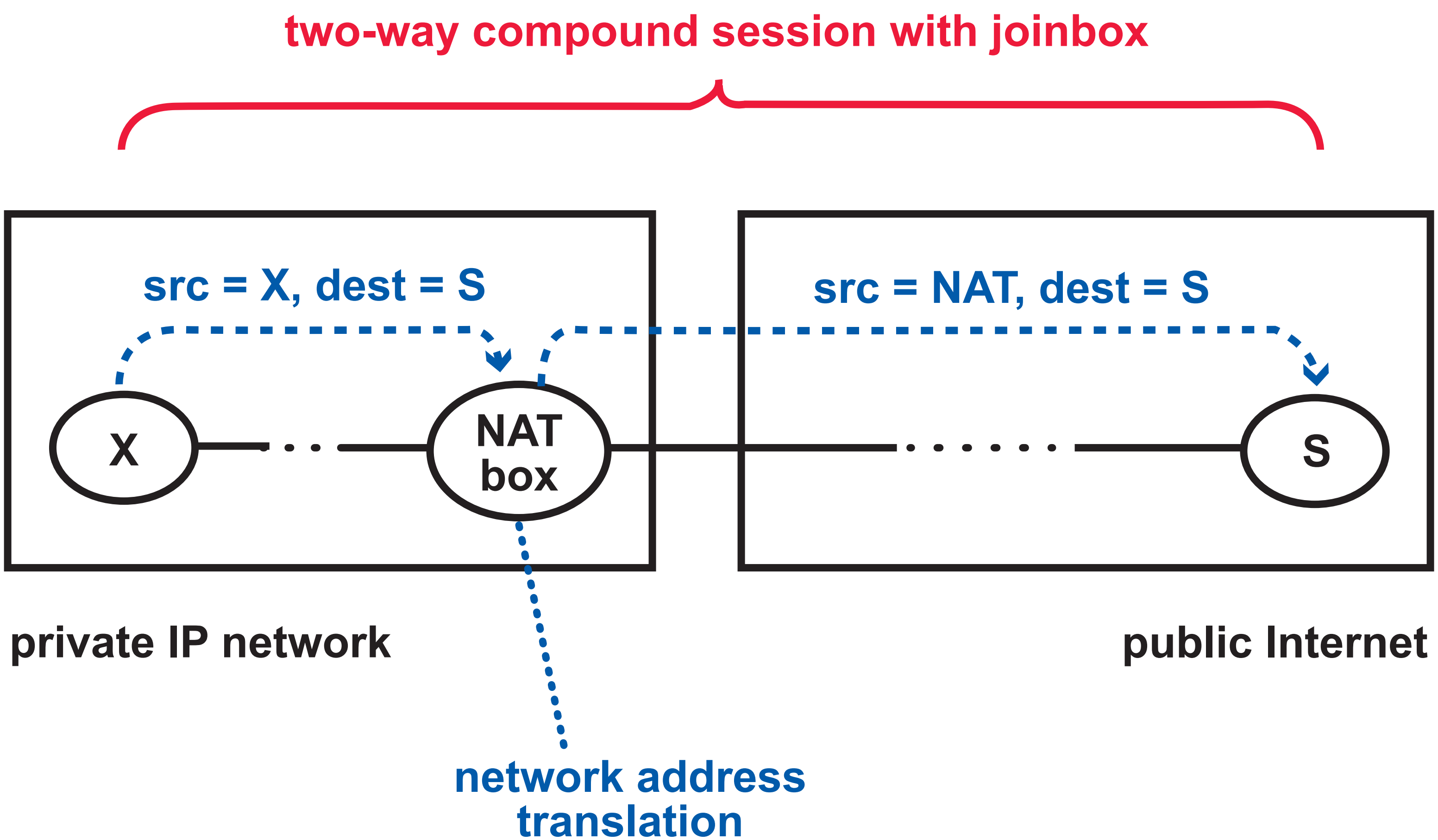


so members of private networks cannot reach each other

IN ADDITION TO COMPOSITION OF NETWORKS . . .

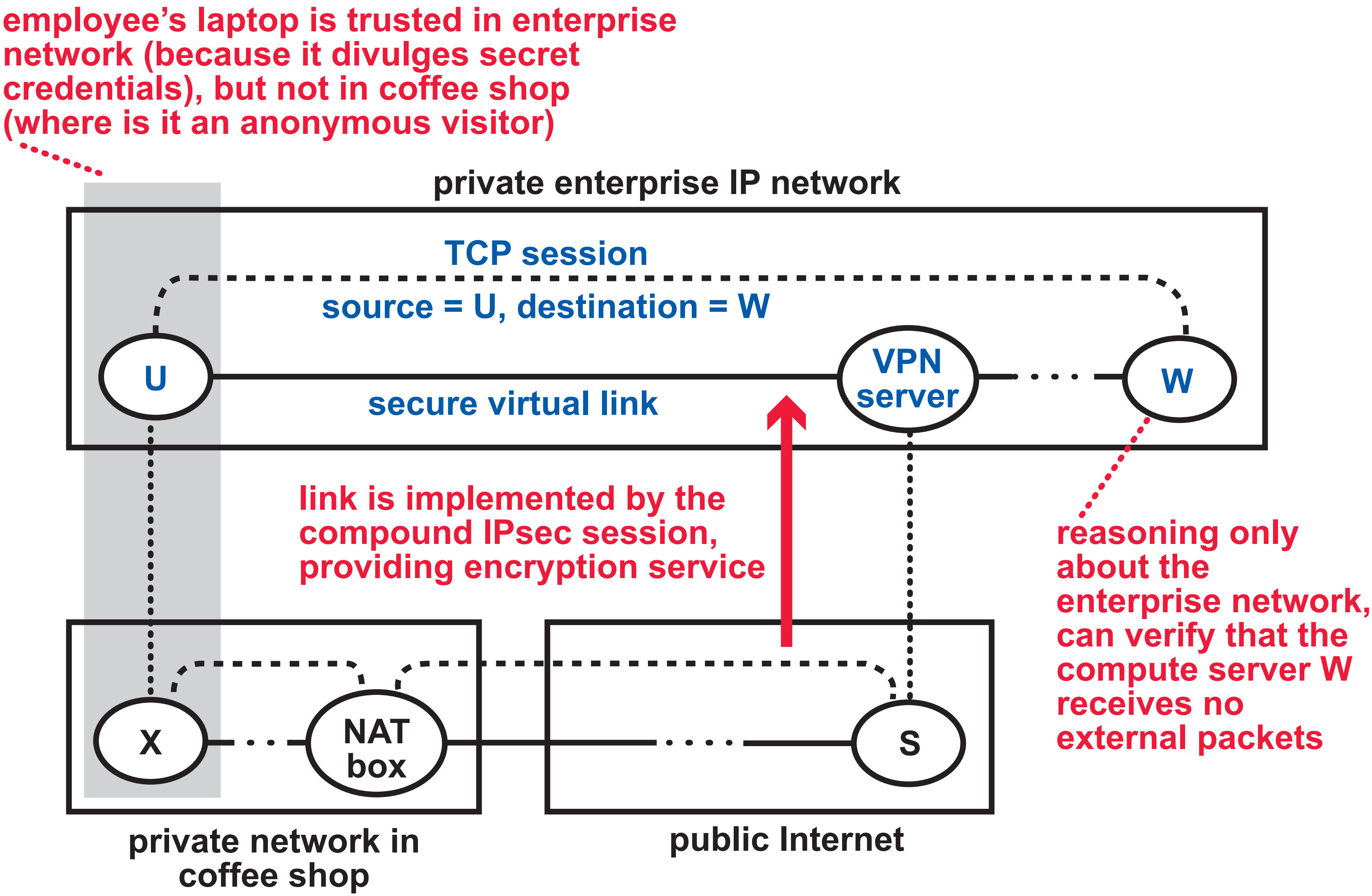
THERE IS SESSION COMPOSITION, . . .

WHICH IS ALSO NOT RECOGNIZED IN THE CLASSIC ARCHITECTURE



BRIDGING, COMPOUND SESSIONS, AND LAYERING:

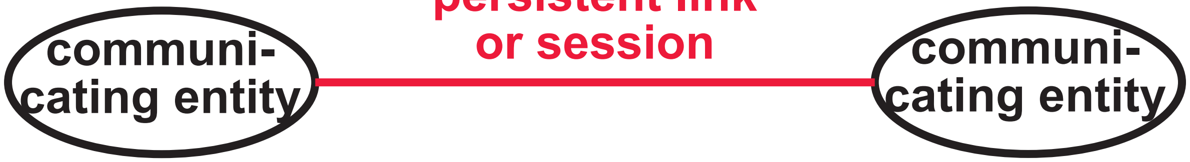
VIRTUAL PRIVATE NETWORKS



A DEFINITION OF MOBILITY

either endpoint can initiate the channel, provided that both are connected to the network

persistent link or session



the channel persists even when one or both endpoints changes its connection to the network

common example: cellphone voice service

persistent link supporting periodic monitoring

wearable health-monitoring device

personalized data analysis and abnormality alerting

this is a data service—
no application programming needed

down the protocol stack

person moves around

device uses both cellular and WiFi connections, alternatively or simultaneously

minimal keep-alive signaling, to reduce battery drain

down the protocol stack

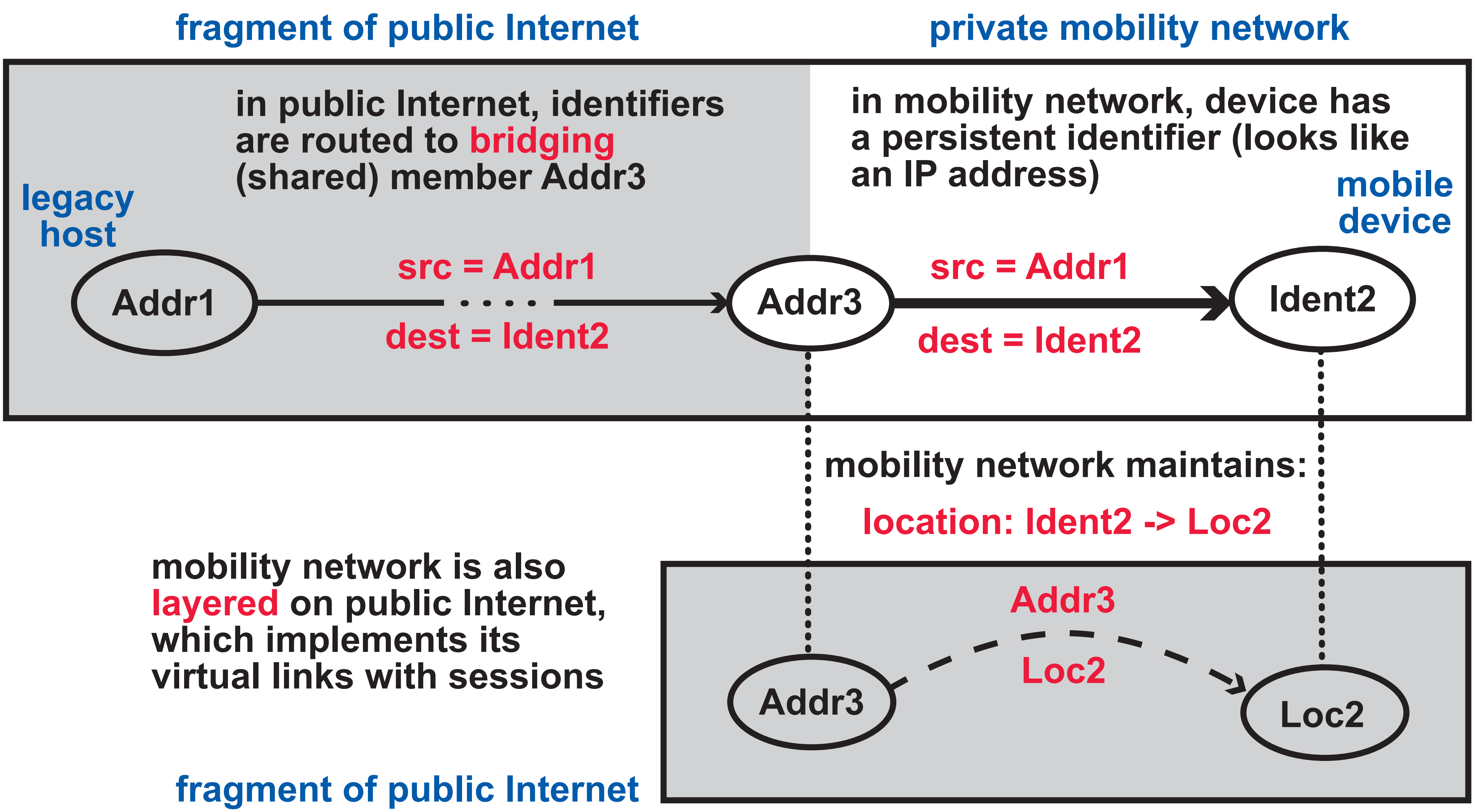
virtual-machine migration

re-routing around failed links to data center

COMPOSITION NOT IN CLASSIC ARCHITECTURE:

LISP-MN FOR MOBILITY

true mobility: a member has a persistent name by which it can be reached at any time, even if it moves during a session

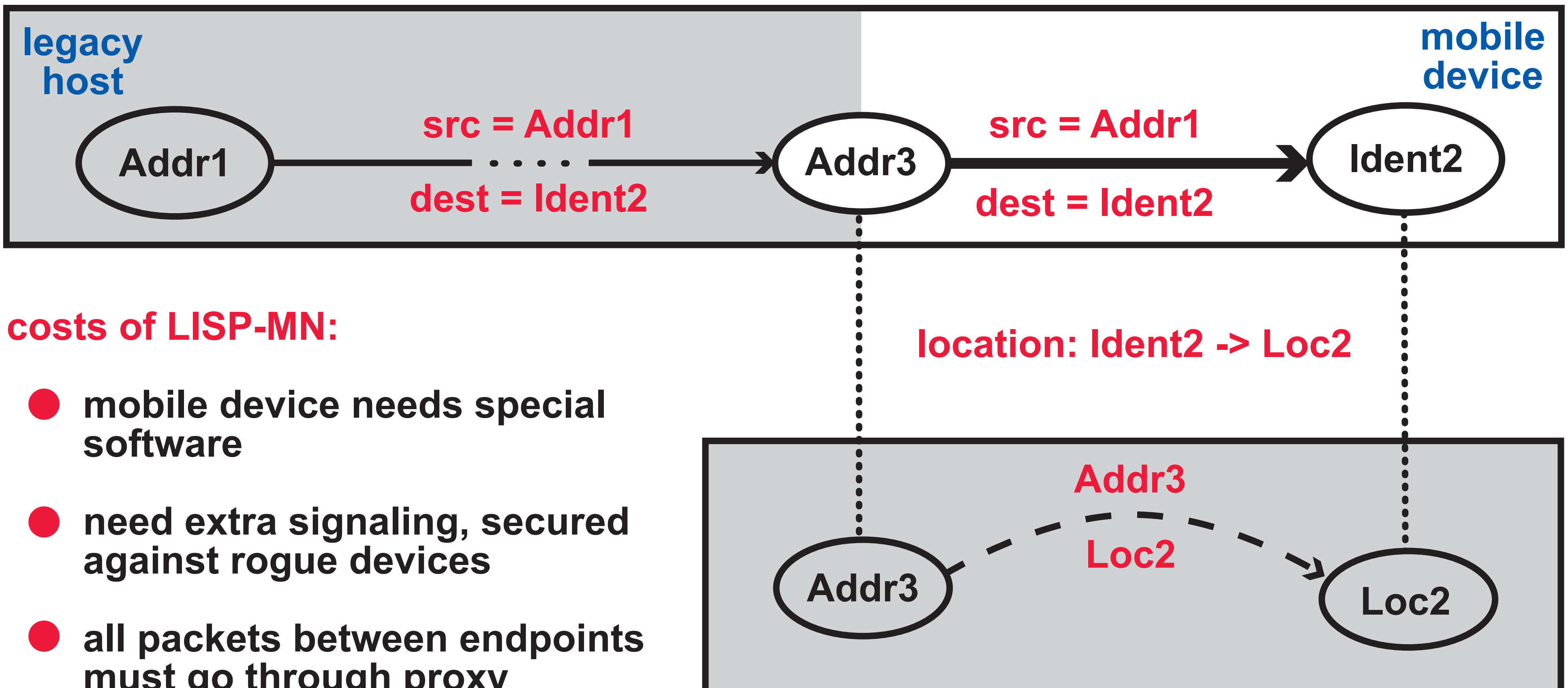


WHY? CAPABILITY IS VERY DIFFICULT TO IMPLEMENT IN THE CLASSIC INTERNET ARCHITECTURE

IP addresses are location-dependent and aggregated for efficient global routing . . .

. . . native Internet mobility would require millions of global routing exceptions, frequently updated

most people get true mobility from cellular service, which is expensive because it does implement dynamic routing to individual devices

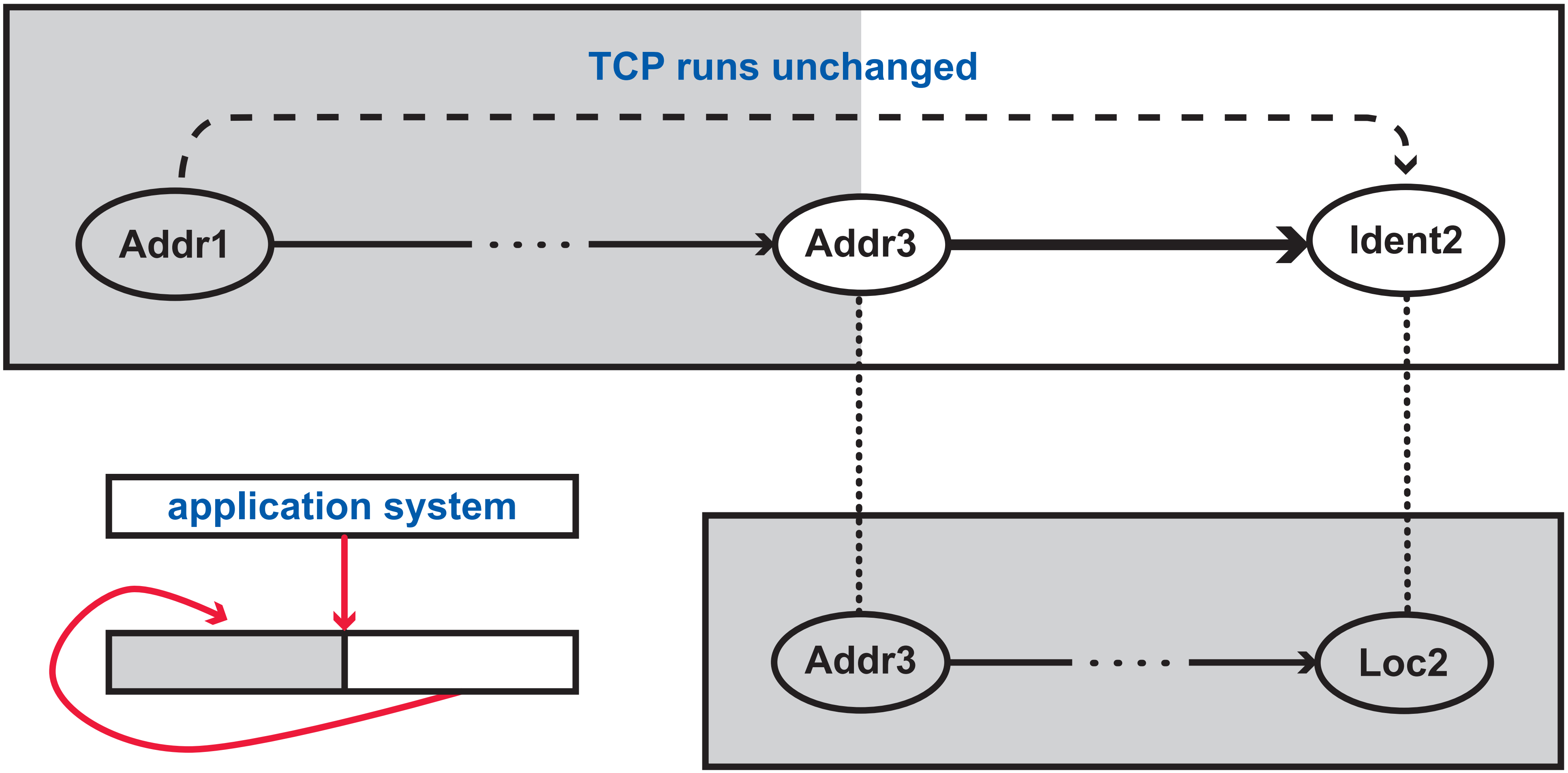


costs of LISP-MN:

- mobile device needs special software
- need extra signaling, secured against rogue devices
- all packets between endpoints must go through proxy

THIS IS A COMMON PATTERN FOR INTEROPERATION OF SPECIAL NETWORKS WITH THE PUBLIC INTERNET

the "observable Internet" is constructed by bridging



although the "usage hierarchy" of networks sometimes has cycles . . .

. . . a dependency graph of links and paths must not have cycles

SUMMARY OF COMPOSITION EXAMPLES

EXAMPLE	WHY IS THERE EXTRA COMPOSITION?	WHAT ABOUT EFFICIENCY?
campus network with VLANs	need two campus-wide views, one for security and one for connectivity, with different topologies	all Ethernets have limited size for efficient broadcast
Virtual Private Network	need a secure network built on top of the public Internet	
LISP-MN for mobility	need a capability that is difficult to implement in the classic Internet architecture	scalable design, with different costs and security vulnerabilities
and many others	Named Data Networking is an experiment with a completely different architecture	SIMPLE makes policy-based routing feasible, by reducing size of forwarding tables

NOW THAT WE HAVE A BETTER MODEL . . .

USE IT TO TEACH NETWORKING

Compositional Network Architecture ·················· graduate-level textbook

- Introduction to Compositional Network Architecture
- Compositional View of the Classic Internet Architecture
- Routing and Forwarding
- Session Protocols
- Middleboxes
- Directories and Mobility
- Network Security
- Ideas for a Better Internet

each chapter describes how an important aspect of networking is realized, across . . .

. . . large and small networks,

. . . general-purpose and special-purpose networks,

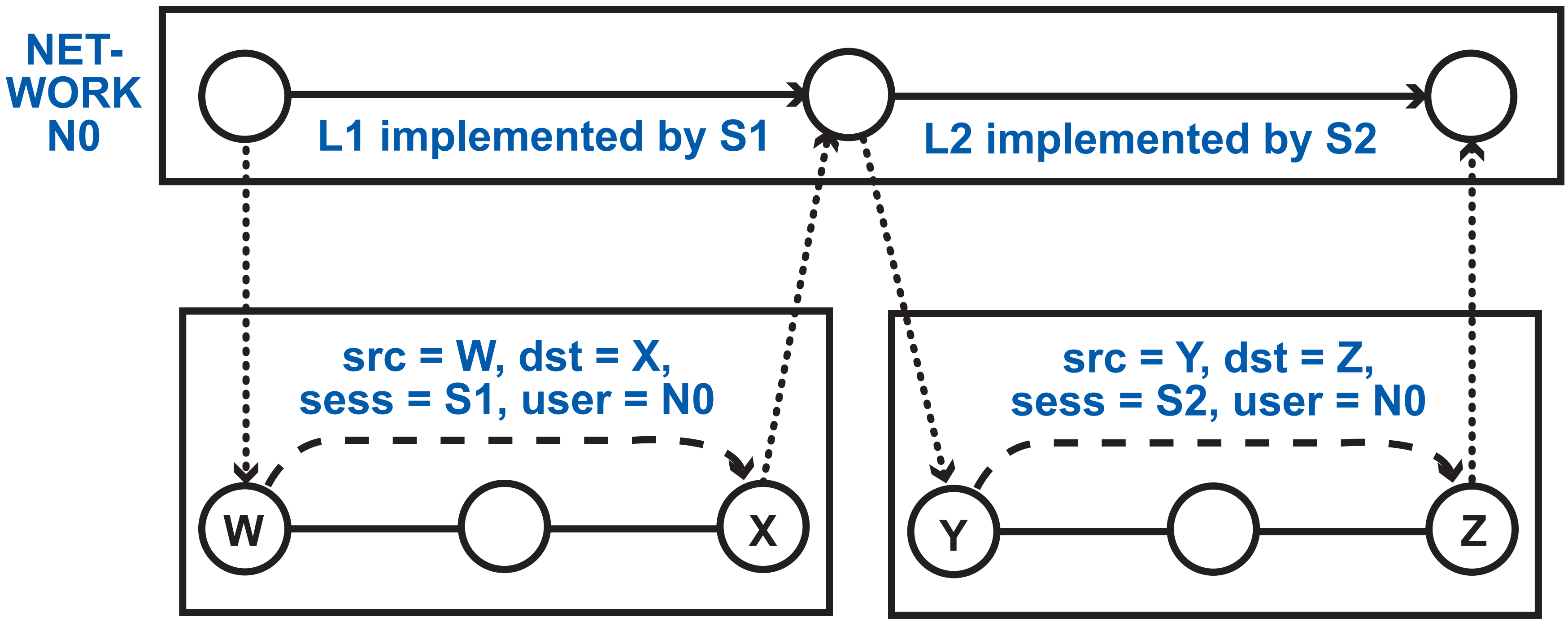
. . . high- and low-level networks

we are emphasizing the *interactions* among these architectural aspects

NOW THAT WE HAVE A BETTER MODEL . . .

USE IT TO PROGRAM NETWORKS

the model gives us re-usable, customizable patterns
to implement (especially for packet processing)



from the patterns
there is a
smooth path to
formalization and
automated analysis

the important optimizations
move functions up (virtualization)
or down (hardware acceleration)
in the network hierarchy . . .

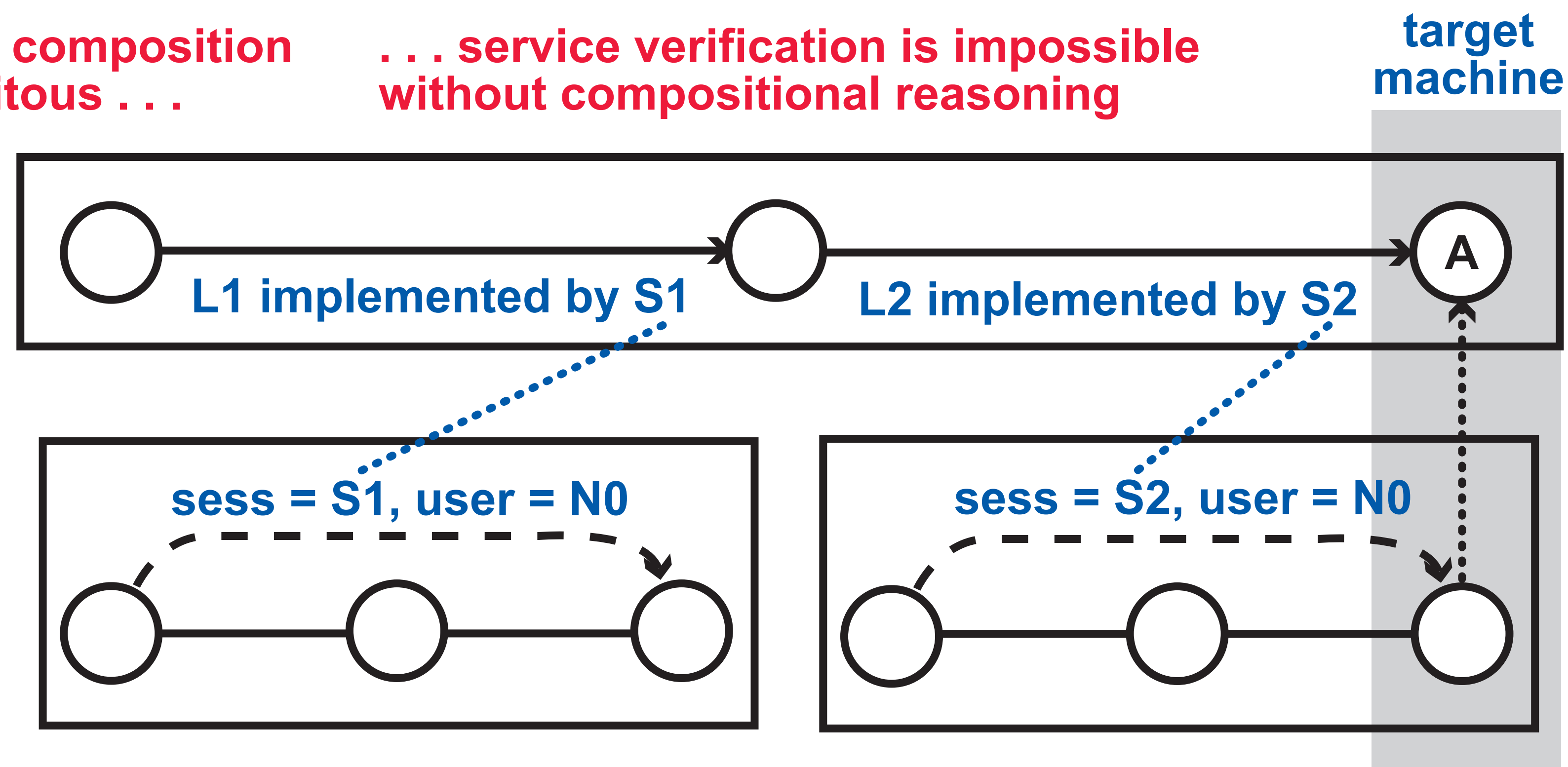
. . . and these can be
automated!

NOW THAT WE HAVE A BETTER MODEL . . .

VERIFICATION OF TRUSTWORTHY SERVICES

because composition is ubiquitous . . .

. . . service verification is impossible without compositional reasoning



with programs derived from the model, it should be possible to verify properties such as “any packet received by target machine is also received by A in N0”

so higher-level networks (with application knowledge) can provide real security for machines—every example uses properties like this

because we can describe security mechanisms precisely in a common model, we are working on a unified proof template for filtering

it includes proofs from opposing sides! (security by filtering versus evasion of censorship)

NOW THAT WE HAVE A BETTER MODEL . . .

USE IT TO UNDERSTAND INTERNET EVOLUTION

COMPOSITION ALLOWS THE CLASSIC INTERNET ARCHITECTURE TO . . .

- interoperate with new concepts
- evolve toward the successful ones

SMOOTHER COMPOSITION WILL MAKE THE PROCESS EASIER AND SAFER

- get rid of unnecessary impediments

IN THE LONG TERM . . .

- What is the optimal way to combine capabilities for network services, e.g., mobility, middleboxes, multihoming, group names, security, enhanced session protocols, etc.?
- What is the best way to satisfy requirements for truly specialized networks, without losing the performance benefits of global best-effort service?

CONCLUSION

the model really matters

COMPOSITIONAL NETWORK ARCHITECTURE IS A PRECISE AND COMPREHENSIVE MODEL FOR DESCRIBING TODAY'S NETWORKS

WHAT ABOUT FORMALIZATION?

- we can't just charge ahead and formalize all of networking
- as always, we must be clever about formalizing pieces that we really need for analysis and verification
- the informal model will help us make sure that the pieces fit together

WHAT ABOUT HAVING AN IMPACT?

- there is not much hope for holistic verification of today's services

too much implementation mess

- we must exploit programmability!

make new implementations that embody the model