Computer Networks

Sándor Laki

ELTE-Ericsson Communication Networks Laboratory

ELTE FI – Department Of Information Systems

lakis@elte.hu

http://lakis.web.elte.hu

Eötvös Loránd University

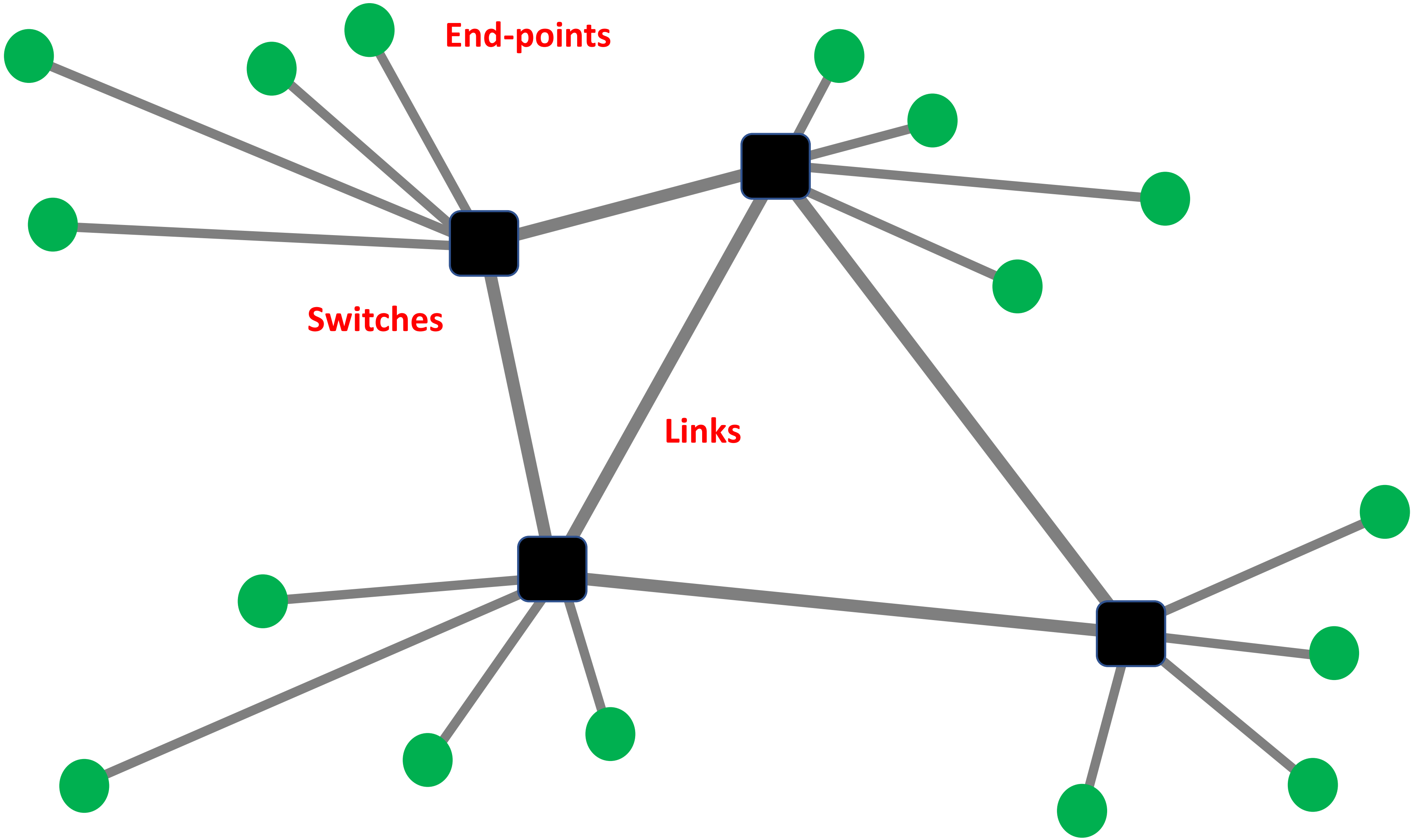
Based on the slides of Laurent Vanbever.
Further inspiration: Scott Shenker & Jennifer Rexford & Phillipa Gill

Last week on
Computer Networks

Overview

What is a network made of?

Three main components



Overview

How to share network resources?

Resource handling
Two different approaches for sharing

Reservation

Reserve the needed bandwidth in advance

Flow-level multiplexing

On-demand

Send data when needed

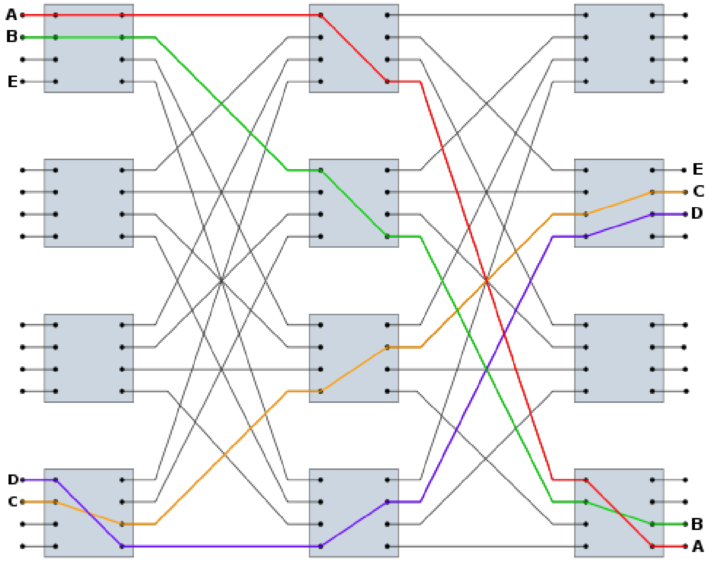
Packet-level multiplexing

Implementation

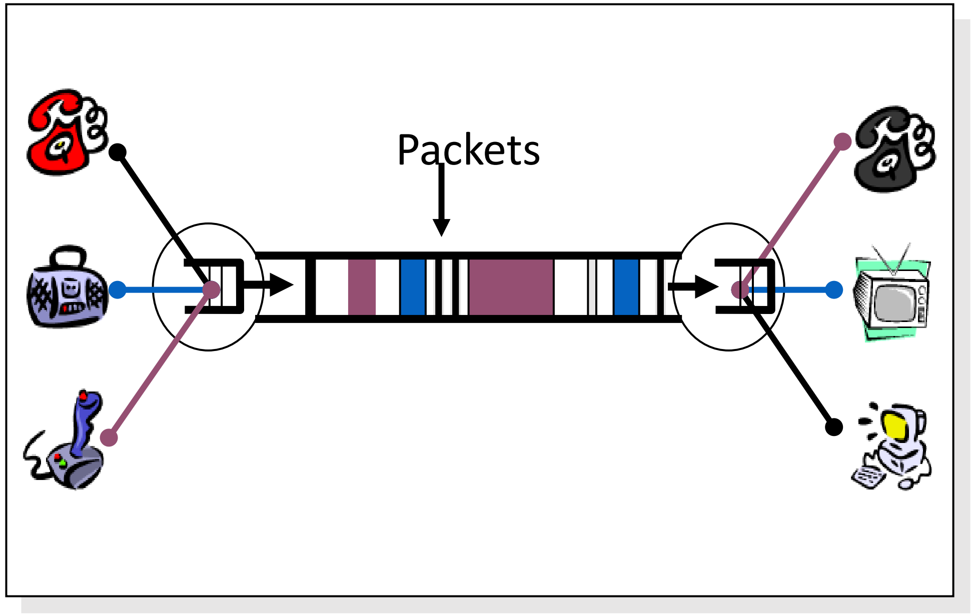
Reservation

On-demand

Circuit-switching
e.g. landline phone networks



Packet-switching
e.g. Internet

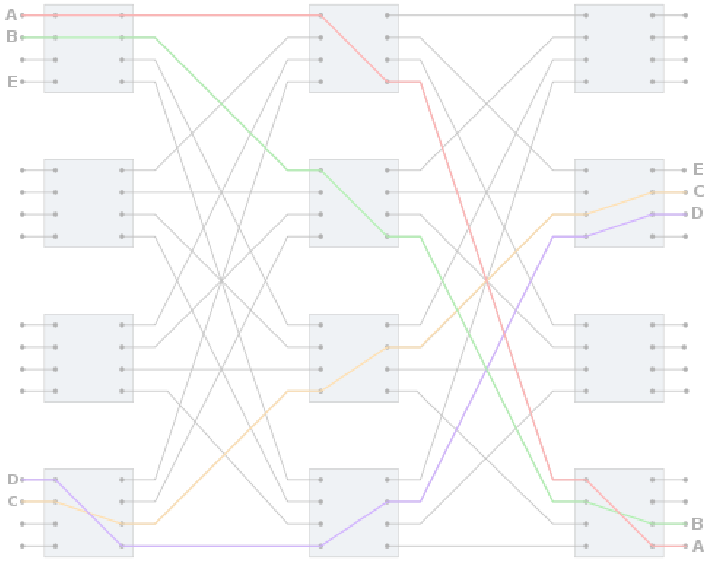


Megvalósítások

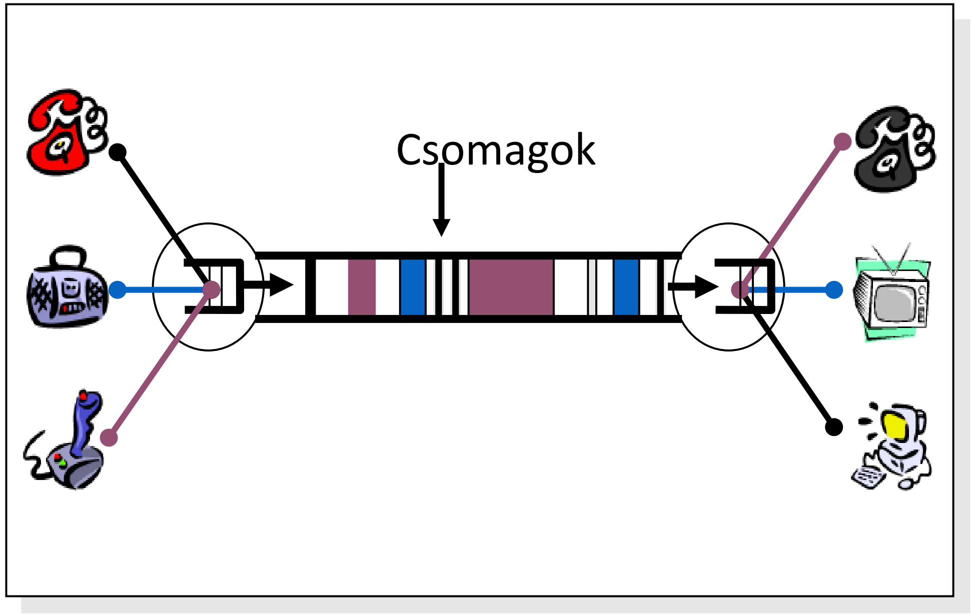
Előre foglalással

Igény szerinti

Áramkörkapcsolt hálózat
Pl. vezetékes telefon

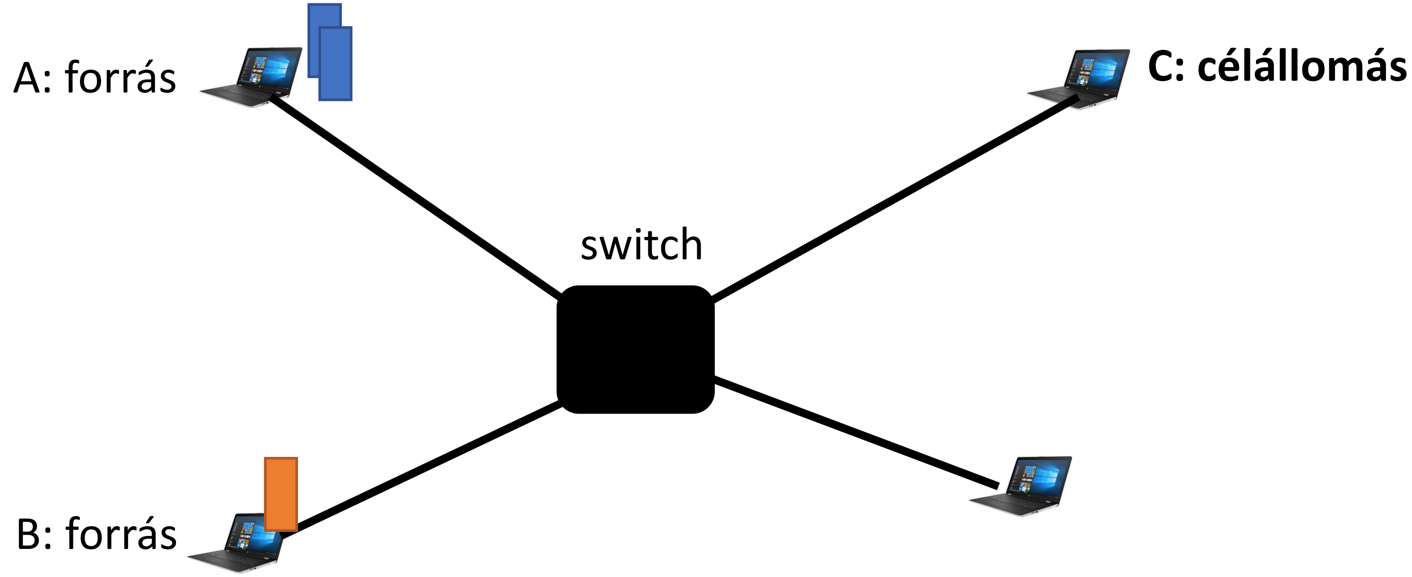


Csomagkapcsolt hálózat
Pl. Internet



Csomagkapcsolt hálózatok

Az adatátvitel egyedi csomagokban történik.



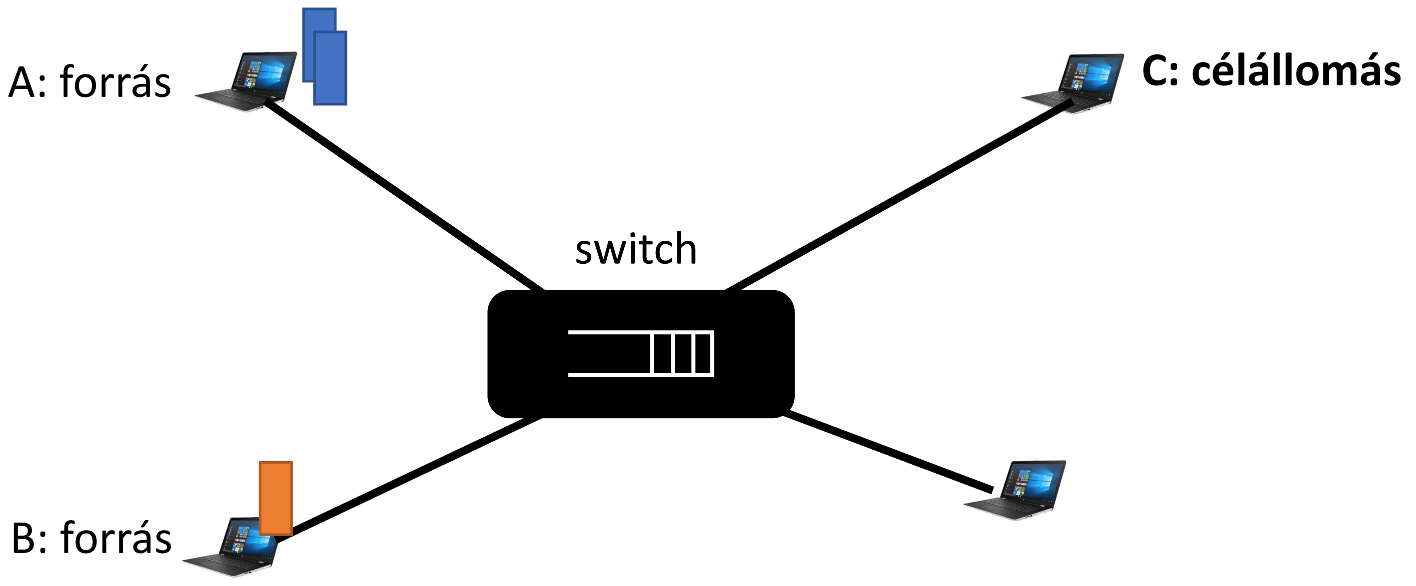
Minden csomag tartamazza a cél címét/azonosítóját (most C).

Nincs globális koordináció, azaz a csomagok zavarhatják egymást. (ld. egyszerre érkeznek be a switchhez)

Pufferelés szükséges a löketek kezeléséhez.

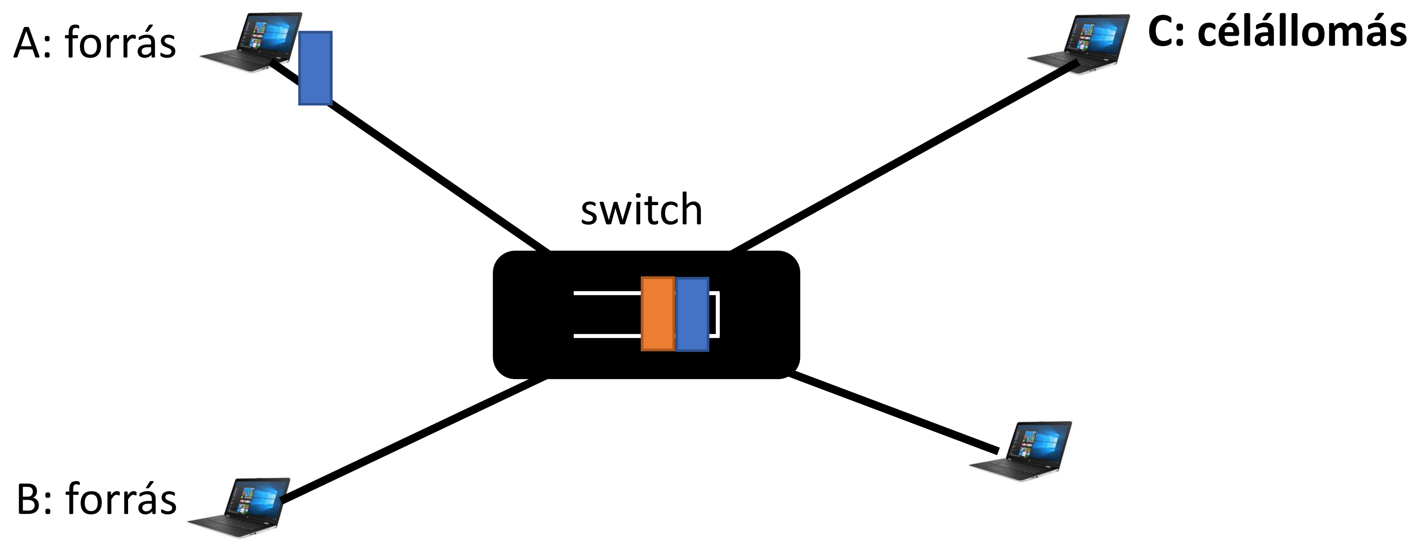
Csomagkapcsolt hálózatok

Pufferelés az átmeneti túlterhelések kezeléséhez

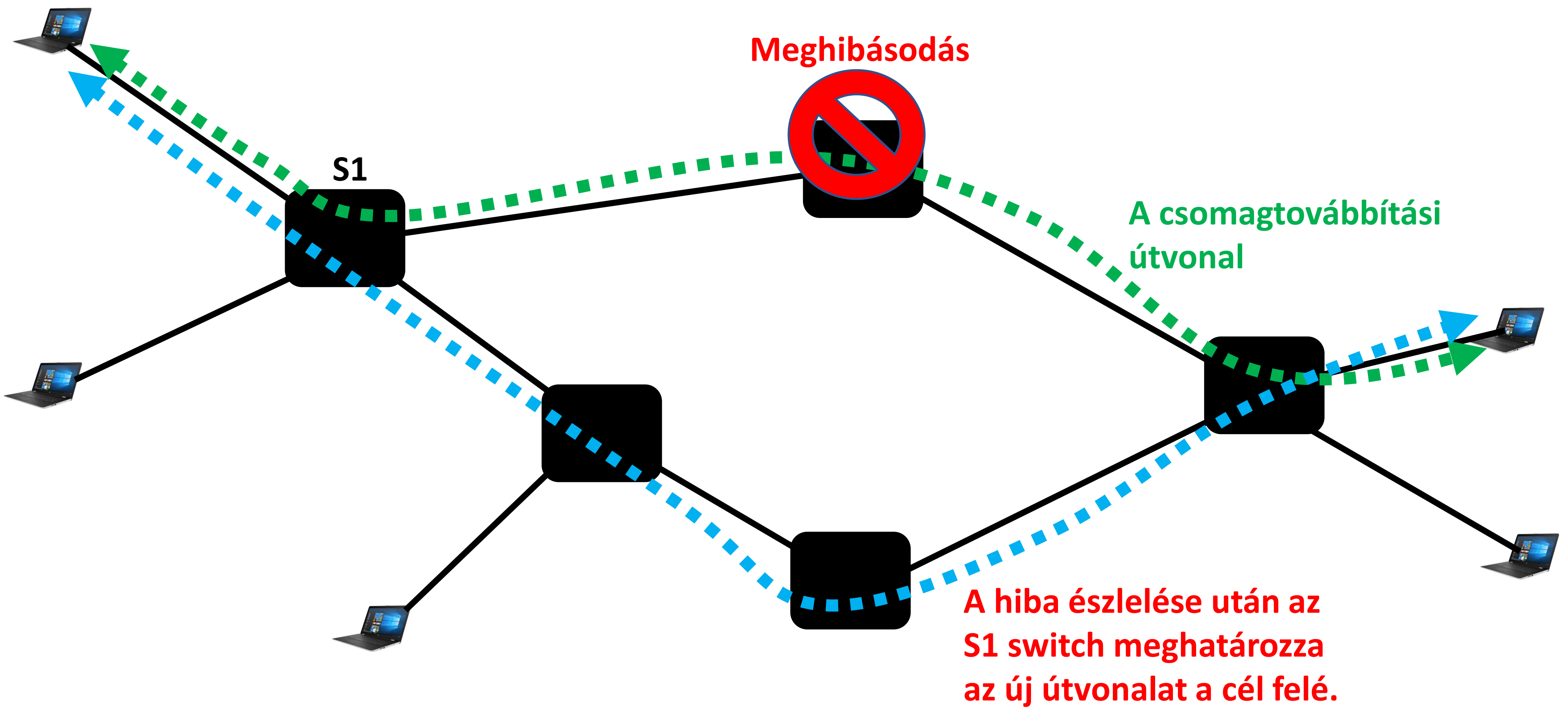


Csomagkapcsolt hálózatok

Pufferelés az átmeneti túlterhelések kezeléséhez



Hiba tolerancia



Érvek ellenérvek

Előnyök

Hatékony erőforrásgazdálkodás

Egyszerű megvalósítás

Hibatolerancia

Hátrányok

Kiszámíthatatlan teljesítmény

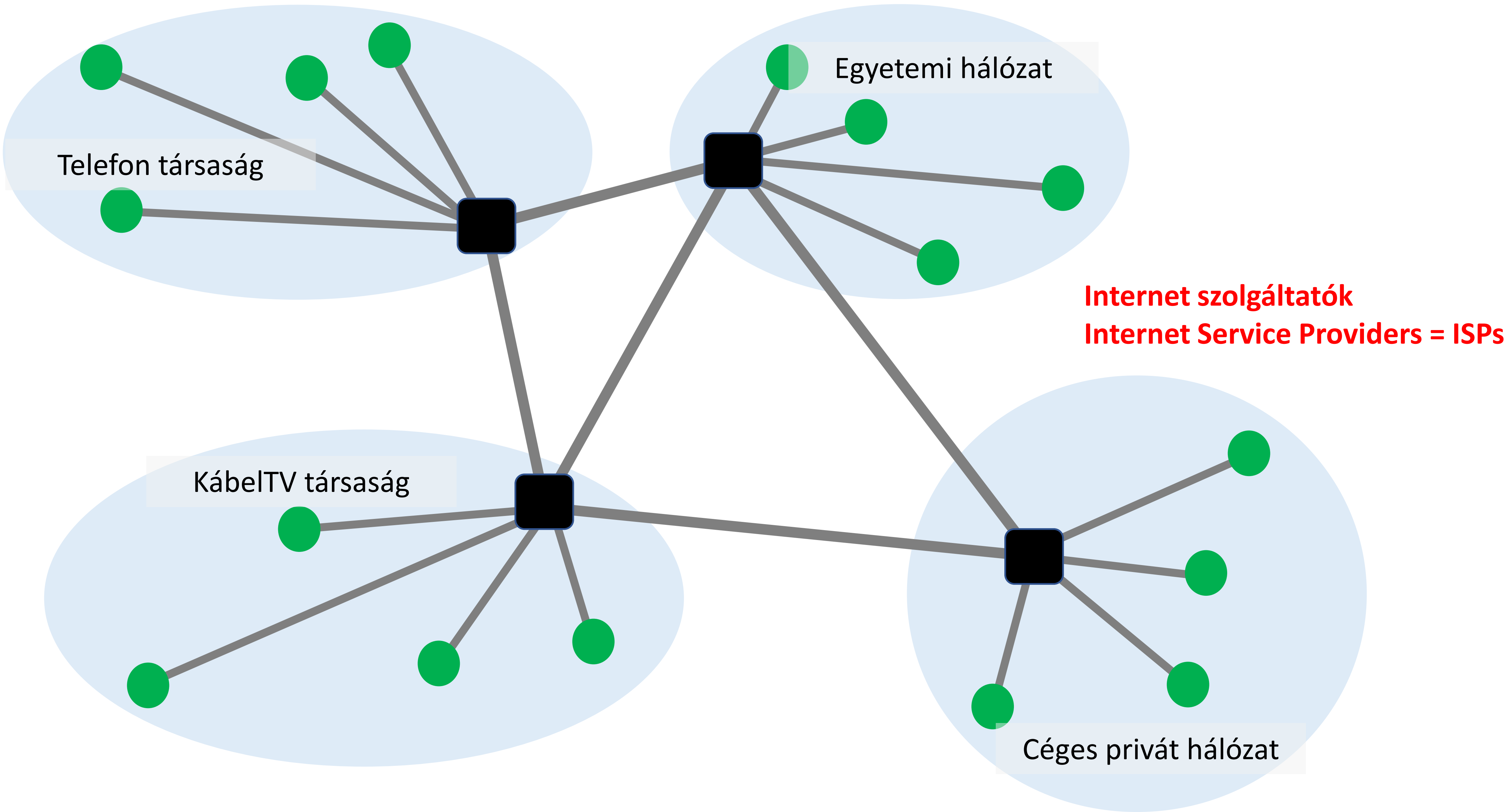
Szükséges puffer-kezelés és torlódás-vezérlés

Az Internet csomagkapcsolt

Rugalmasság és hatékonyság

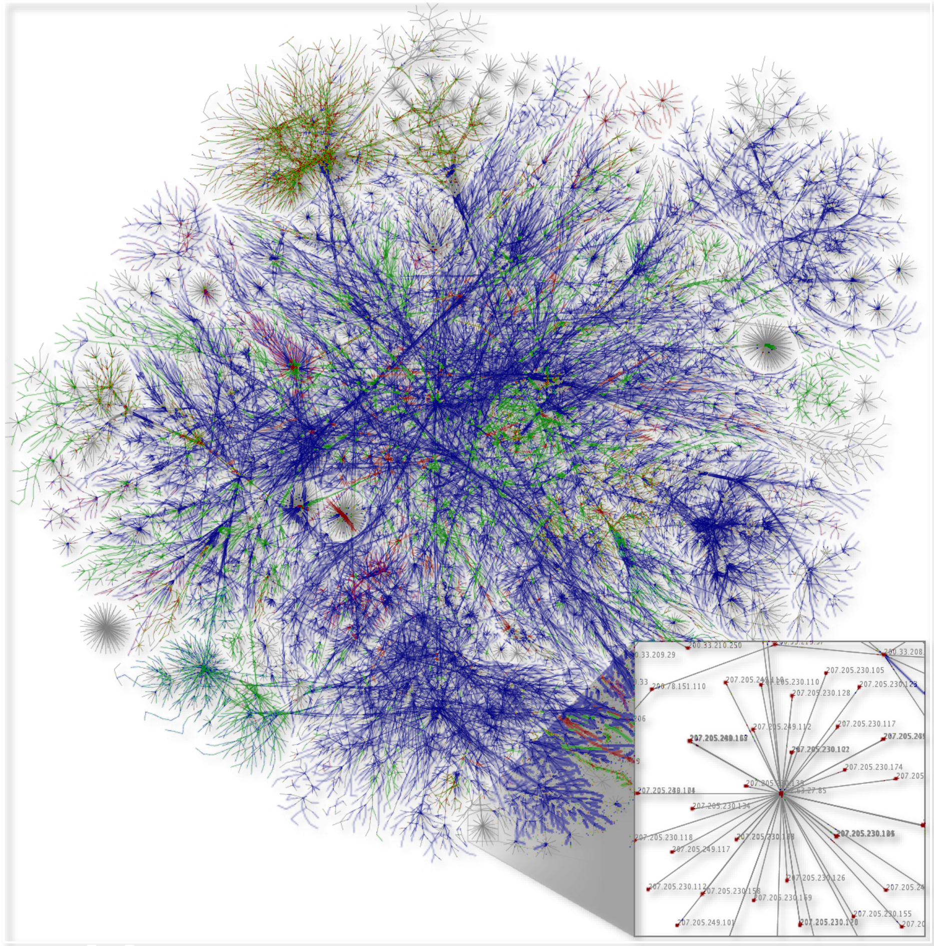
Áttekintés

Hogyan szervezzük a hálózatot?

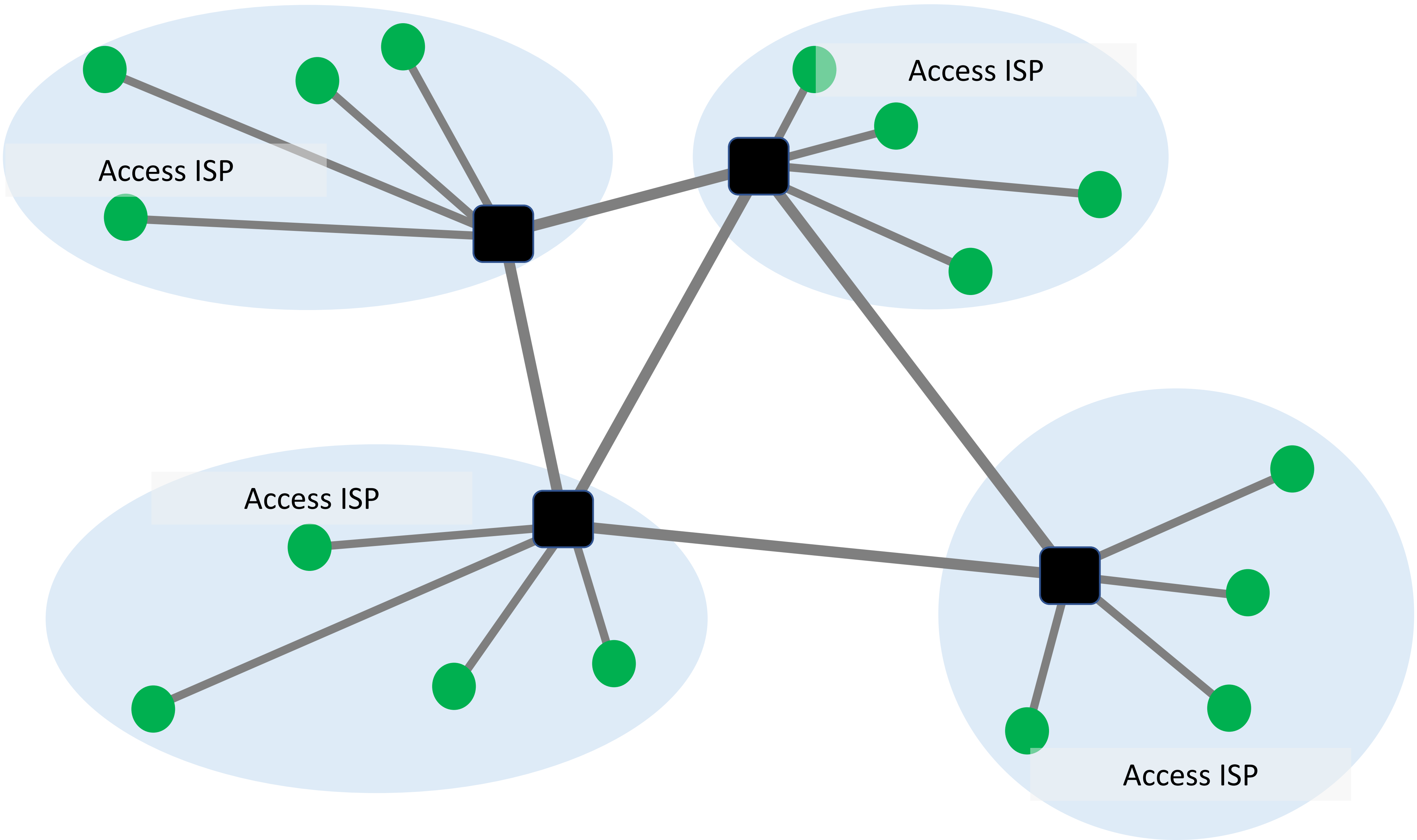

Telefon társaság
Egyetemi hálózat
Internet szolgáltatók
Internet Service Providers = ISPs
KábelTV társaság
Céges privát hálózat

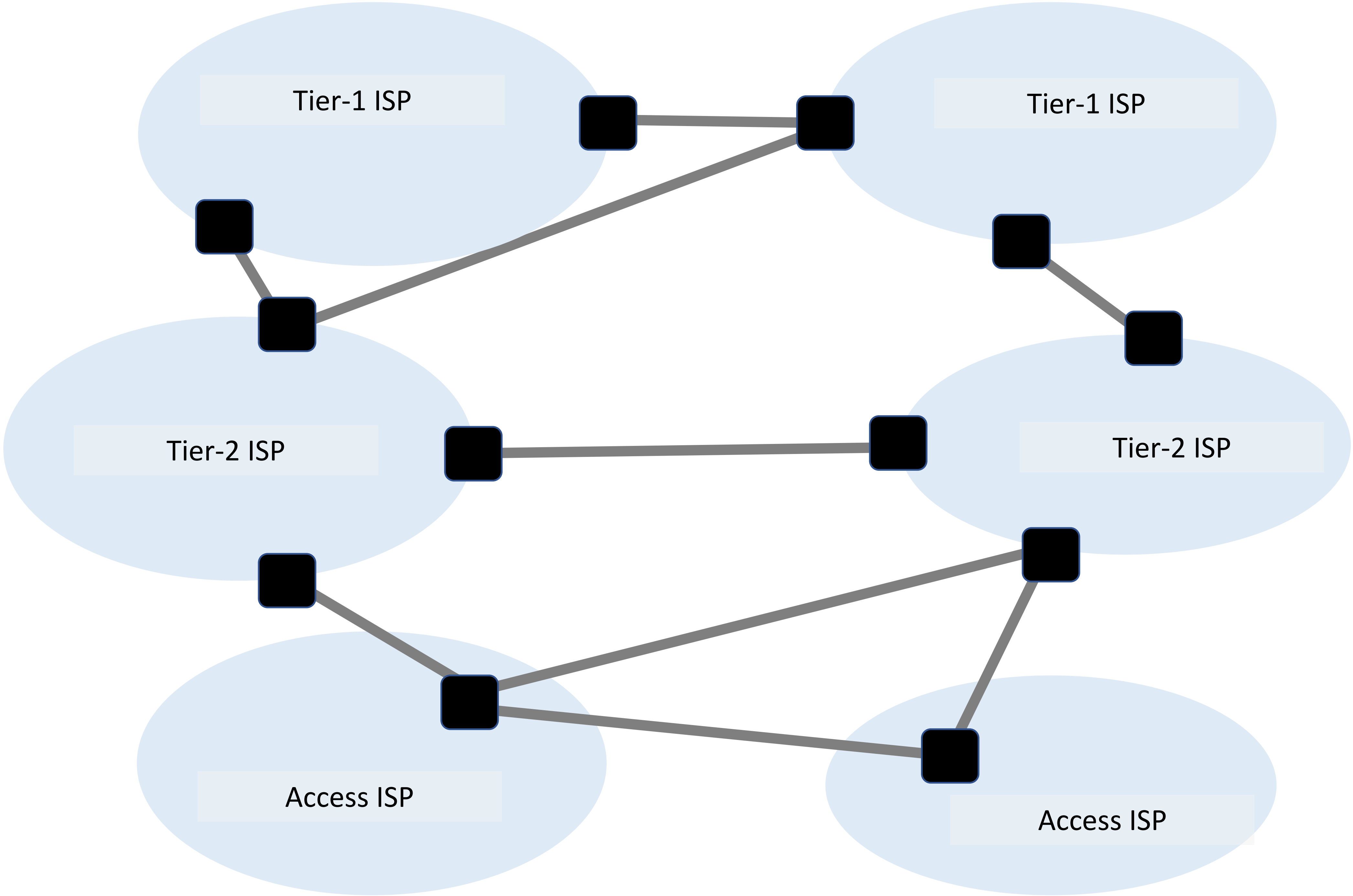
Mi az internet?

- Hálózatok hálózata
- A világra kiterjedő nyitott WAN
- Jellemzői
  - rendszerfüggetlenség;
  - nincs központi felügyelet;
  - építőelemei a LAN-ok;
  - globális;
  - olyan szolgáltatásokat nyújt, mint a **W**orld **W**ide **W**eb, e-mail vagy fájlátvitel.



ISP – Internet szolgáltató

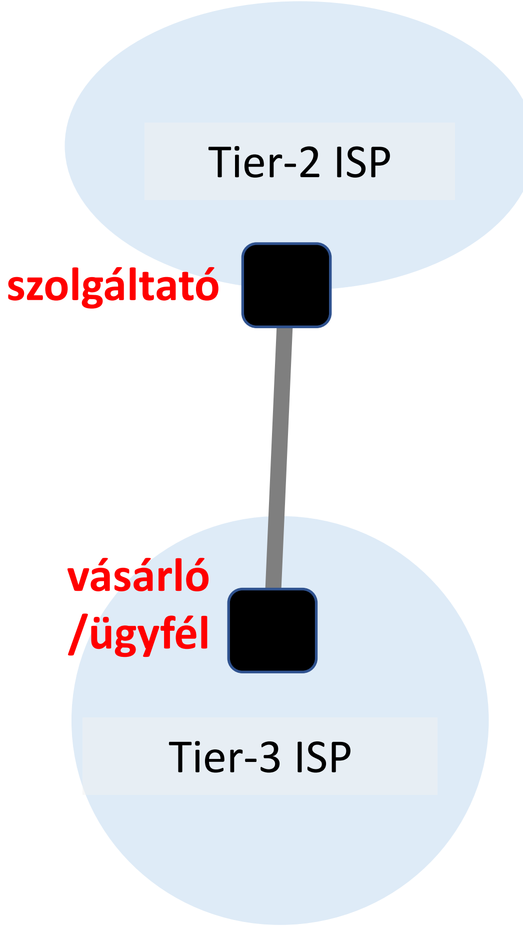


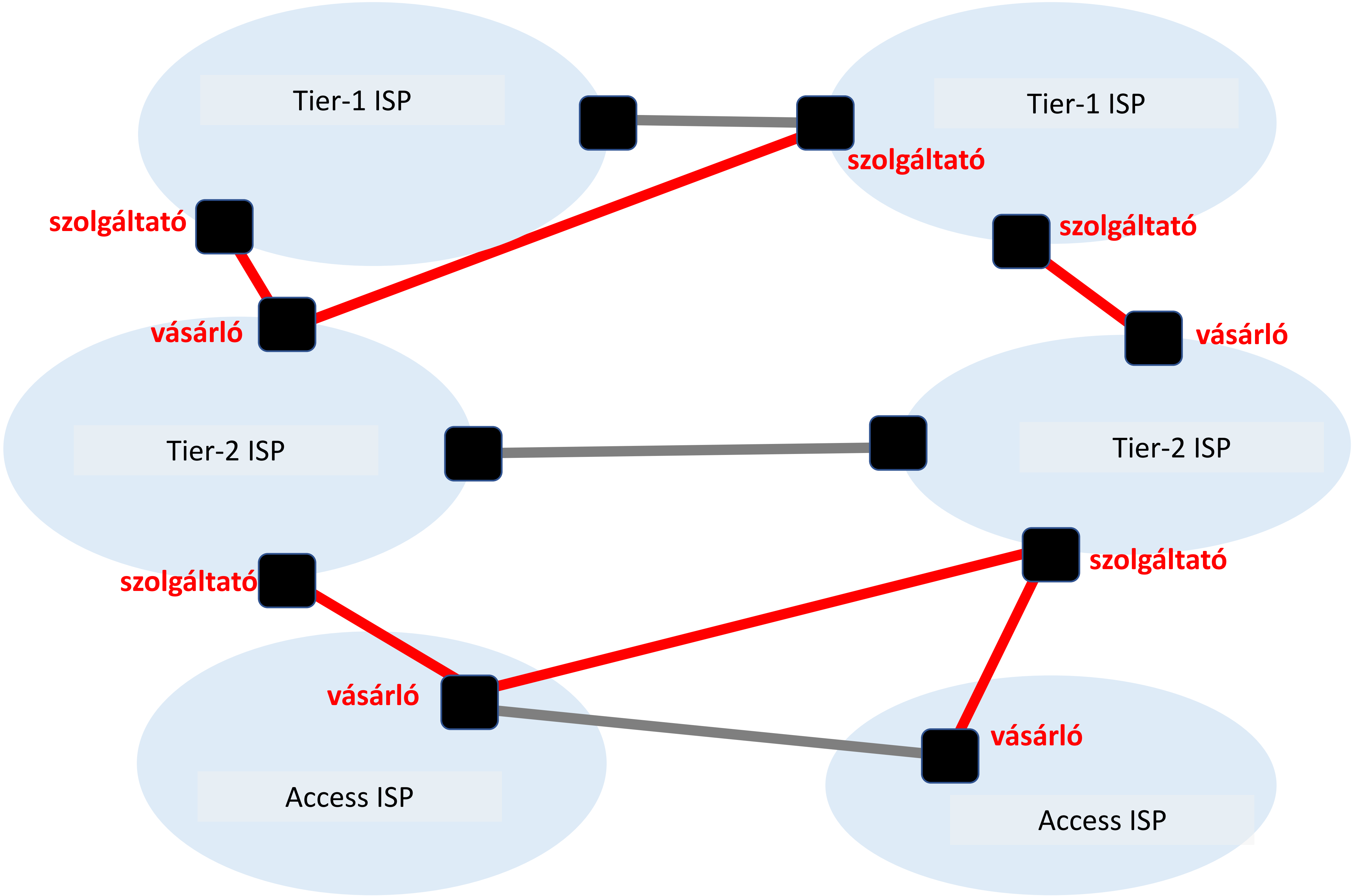

Tier-1 ISP
Tier-1 ISP
Tier-2 ISP
Tier-2 ISP
Access ISP
Access ISP

Az Internet hierarchikus struktúrája

szolgáltató-vásárló (provider-customer) viszonyok

Tier-1 nemzetközi	Nincs szolgáltatója
Tier-2 nemzeti	Tier-3 szolgáltatóknak nyújt átjárást Legalább egy szolgáltatója van
Tier-3 helyi	Nem nyújt átjárást más szolgáltatóknak Legalább egy szolgáltatója van




Tier-1 ISP
Tier-1 ISP
szolgáltató
szolgáltató
szolgáltató
vásárló
vásárló
Tier-2 ISP
Tier-2 ISP
szolgáltató
szolgáltató
vásárló
vásárló
Access ISP
Access ISP

Hálózatok eloszlása a Tier-ekben

~50.000 hálózat összesen

Tier-1
nemzetközi

Nincs szolgáltatója

pár tucat

Tier-2
nemzeti

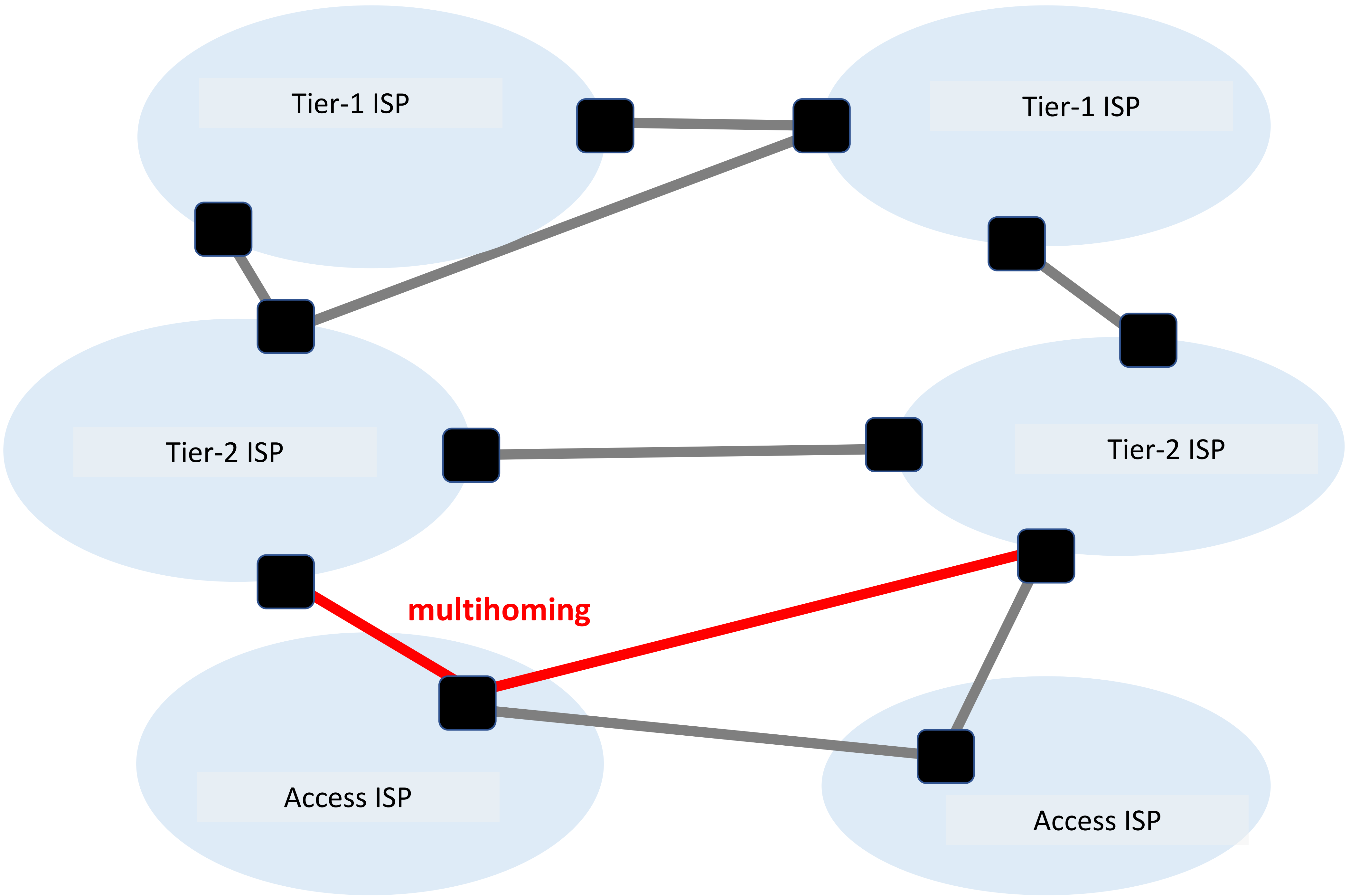
Tier-3 szolgáltatóknak nyújt átjárást
Legalább egy szolgáltatója van

több ezer

Tier-3
helyi

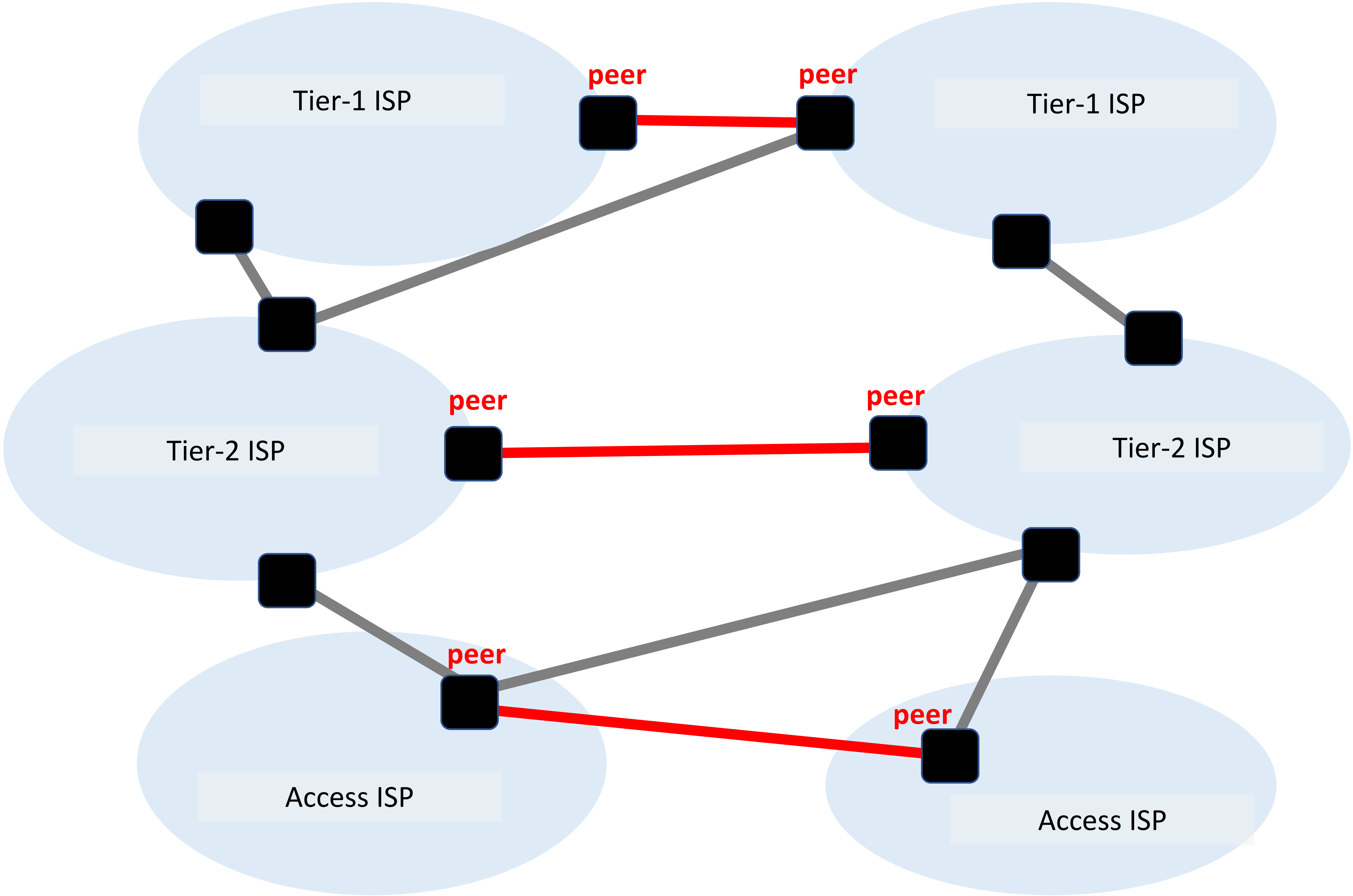
Nem nyújt átjárást más szolgáltatóknak
Legalább egy szolgáltatója van

85-90%


Tier-1 ISP
Tier-1 ISP
Tier-2 ISP
Tier-2 ISP
multihoming
Access ISP
Access ISP

Némely hálózat között közvetlen kapcsolat is létezik – csökkenti a szolgáltatónak fizetendő számlát

Ezt hívjuk „peering”-nek – ez egyfajta kölcsönös kapcsolat…


Tier-1 ISP
peer
peer
Tier-1 ISP
peer
Tier-2 ISP
peer
Tier-2 ISP
peer
peer
Access ISP
Access ISP

A szomszédos hálózatok egyesével való összekapcsolása túl költséges lenne

Infrastruktúra költségek

Fizikai linkek kiépítése vagy bérlése

Sávszélesség költségek

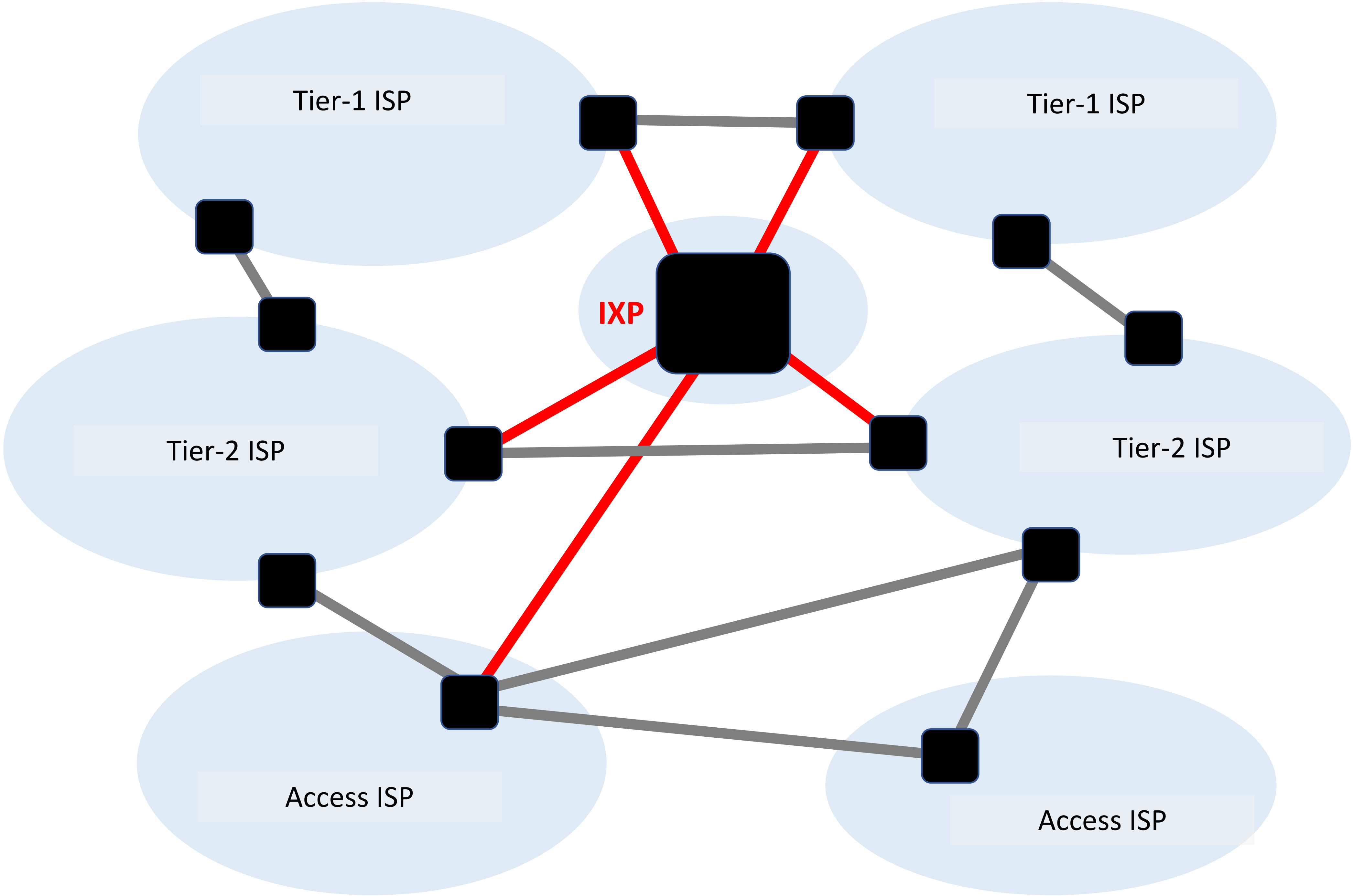
A linkek nem feltétlenül lesznek teljesen kihasználva

Humán költségek

Minden kapcsolatot egyedi módon kell kezelni

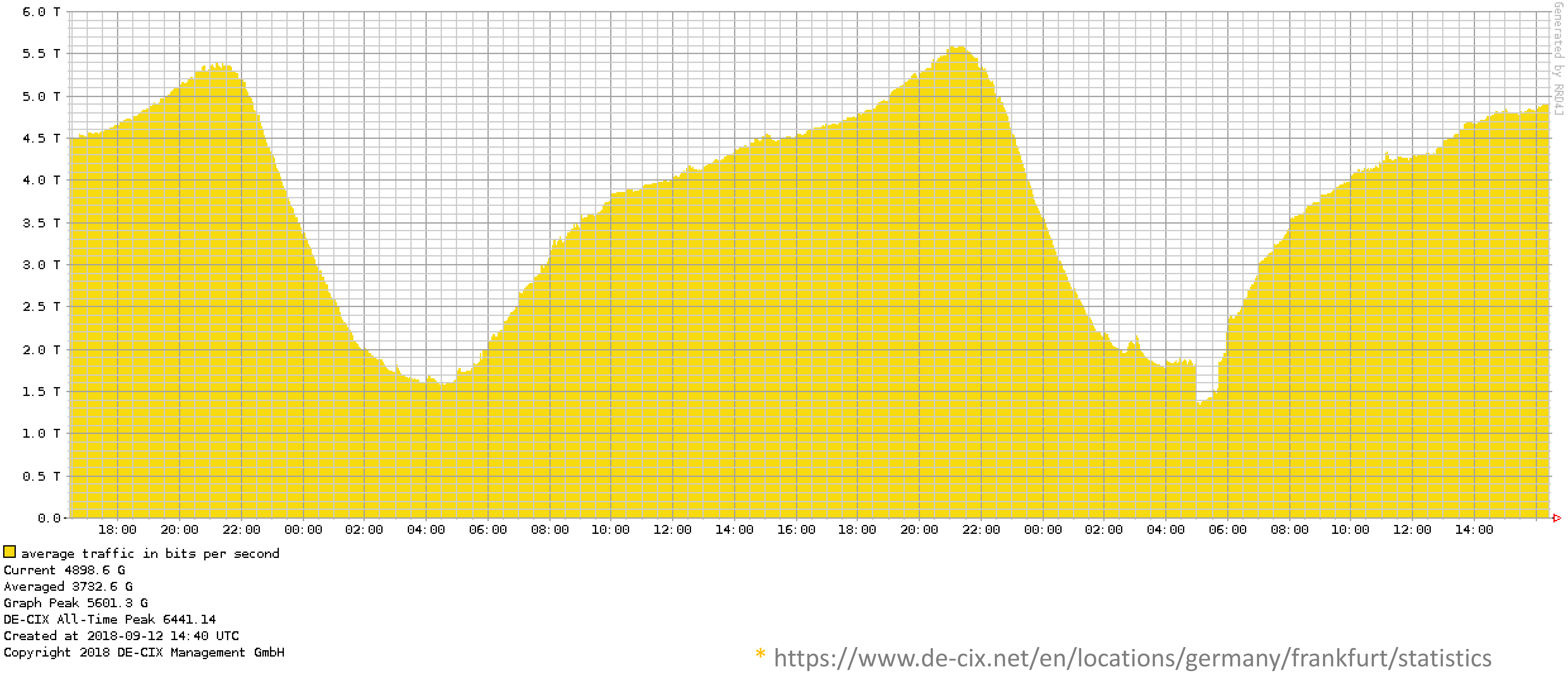
A problémát az úgynevezett Internet eXchage Pontok (IXP) oldják meg

Az IXP-k lehetővé teszik több hálózat összekapcsolását egy fizikai (földrajzi/topológiai) helyen.



Tier-1 ISP
Tier-1 ISP
IXP
Tier-2 ISP
Tier-2 ISP
Access ISP
Access ISP

Egy IXP két napja – DE-CIX Frankfurt



* https://www.de-cix.net/en/locations/germany/frankfurt/statistics

This week

How does communication happen?

How do we characterize it?

Briefly…

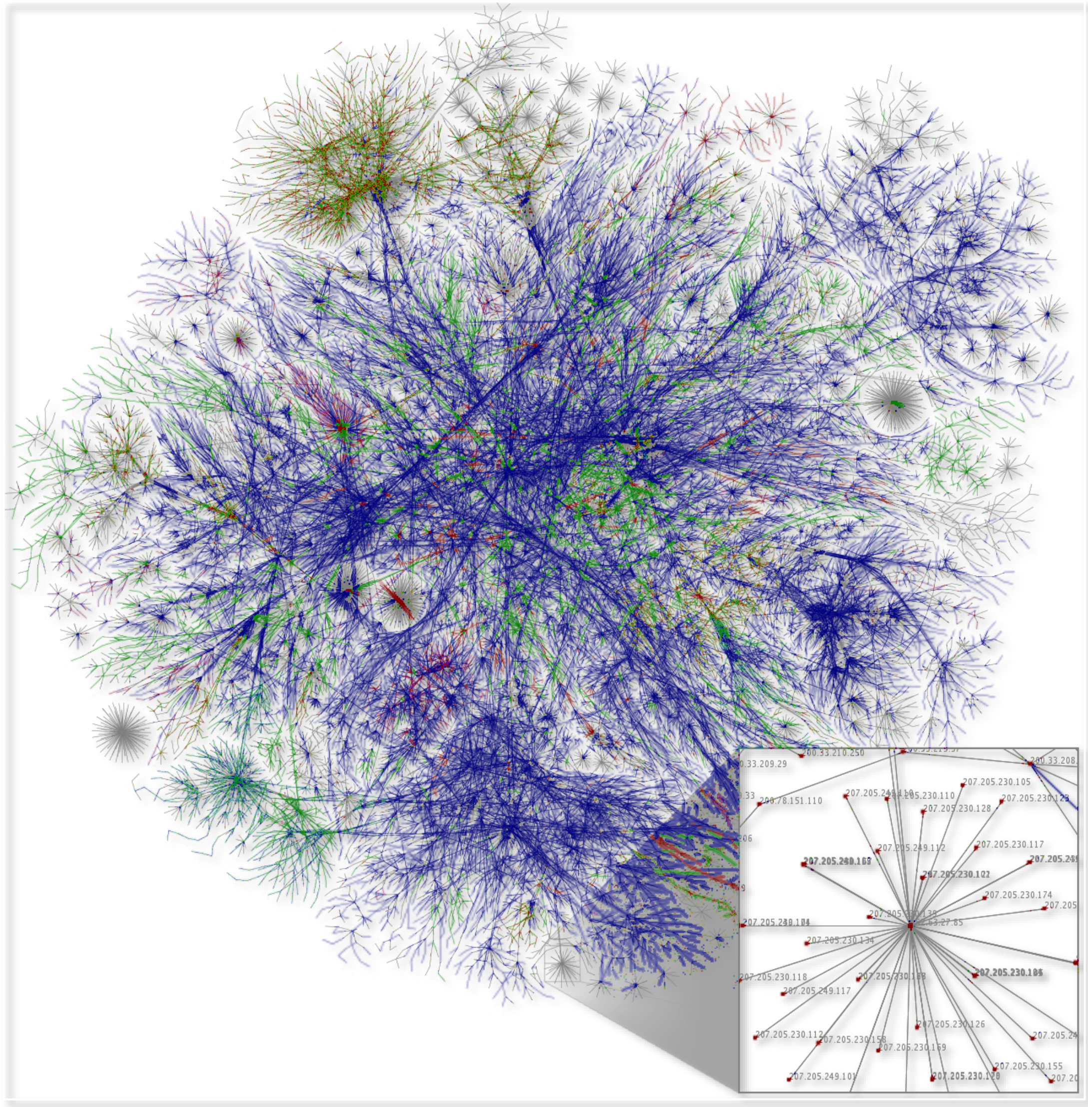
The Internet should allow

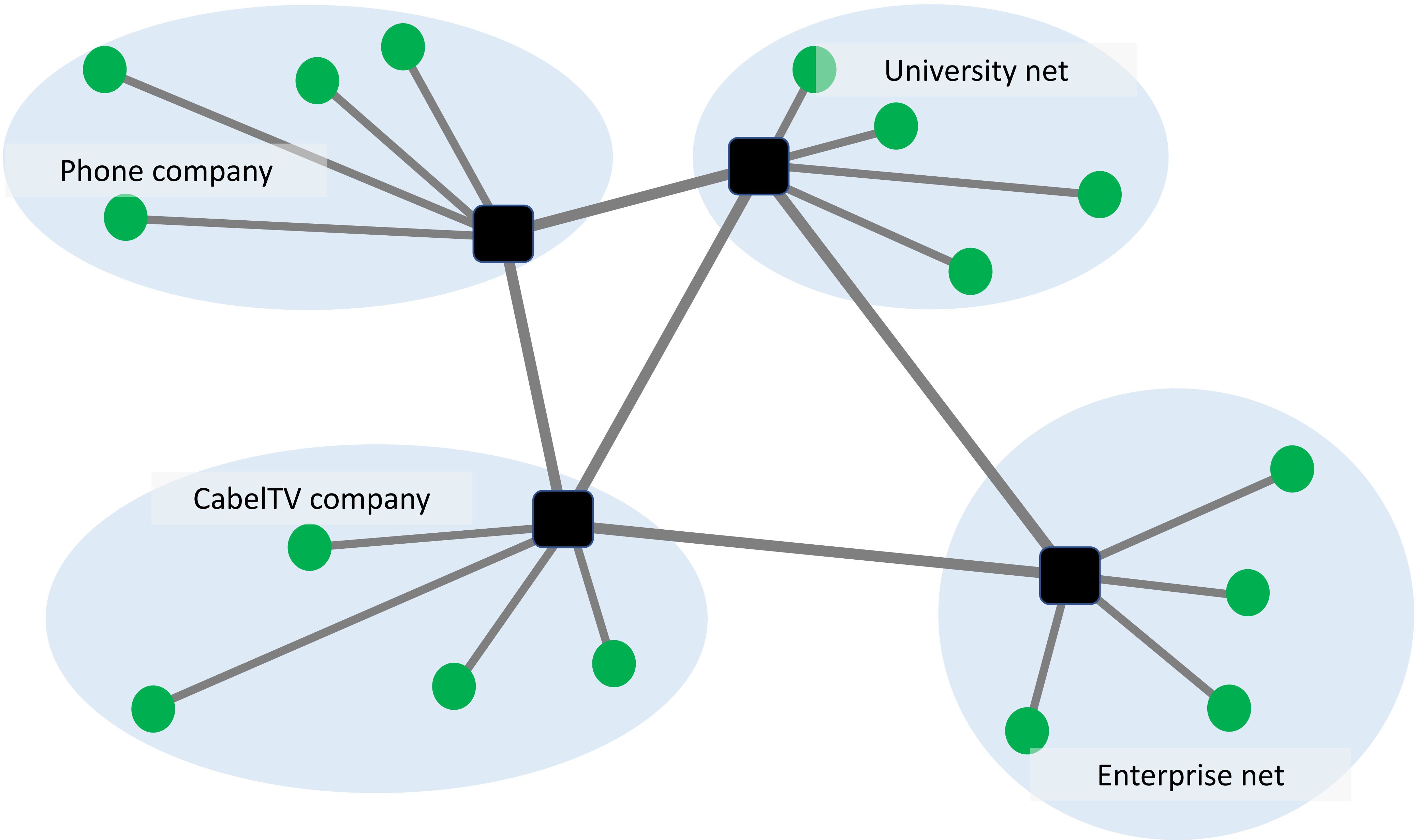
processes on different hosts

to exchange data

everything else is just commentary…

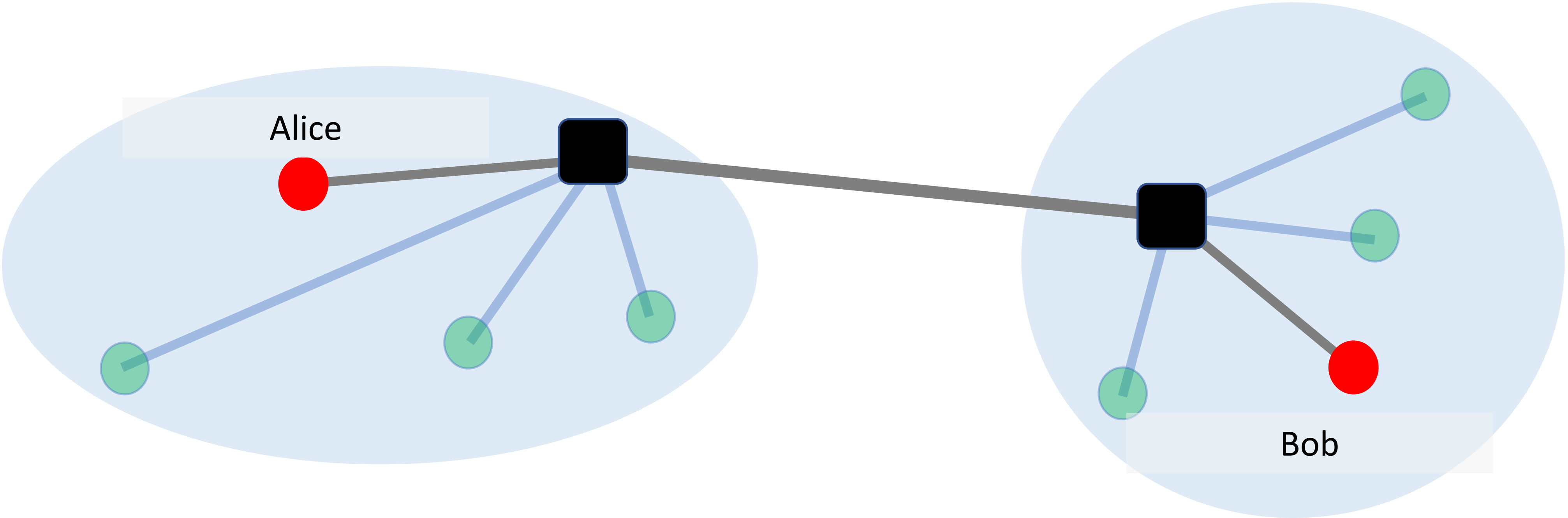
Ok, but how to do that in a complex system like the Internet?





Phone company
University net
CabelTV company
Enterprise net

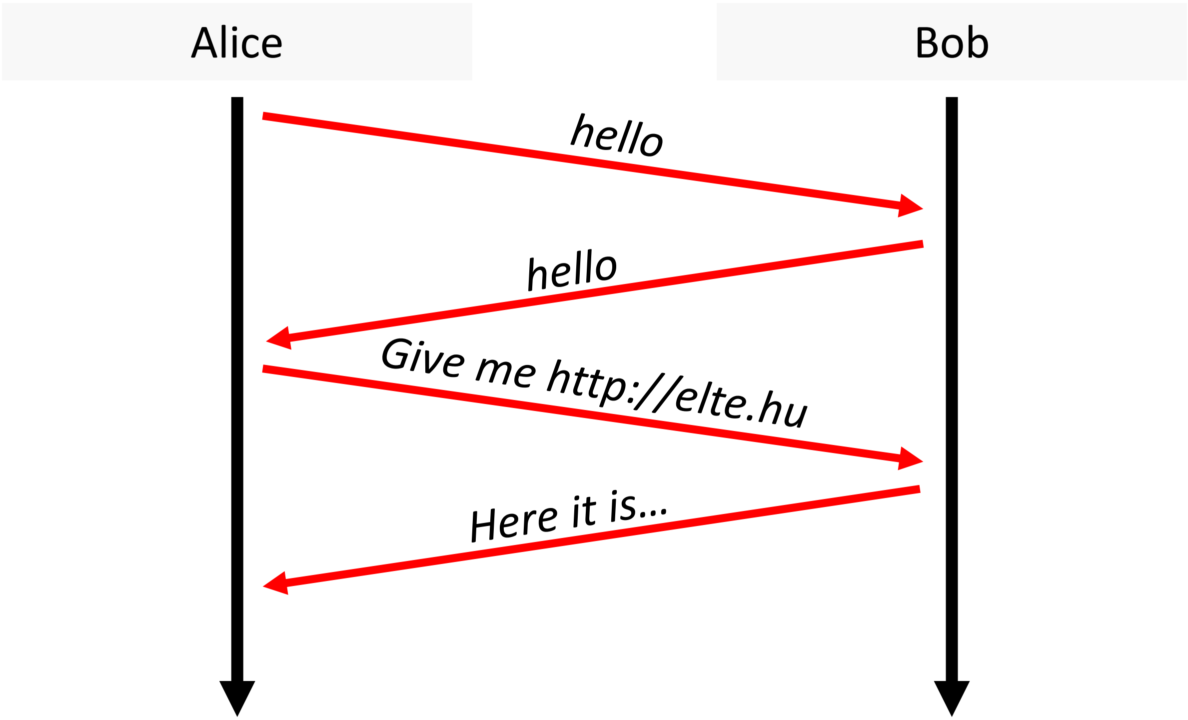
To exchange data, Alice and Bob use
a set of network protocols



A protocol is like a conversational convention

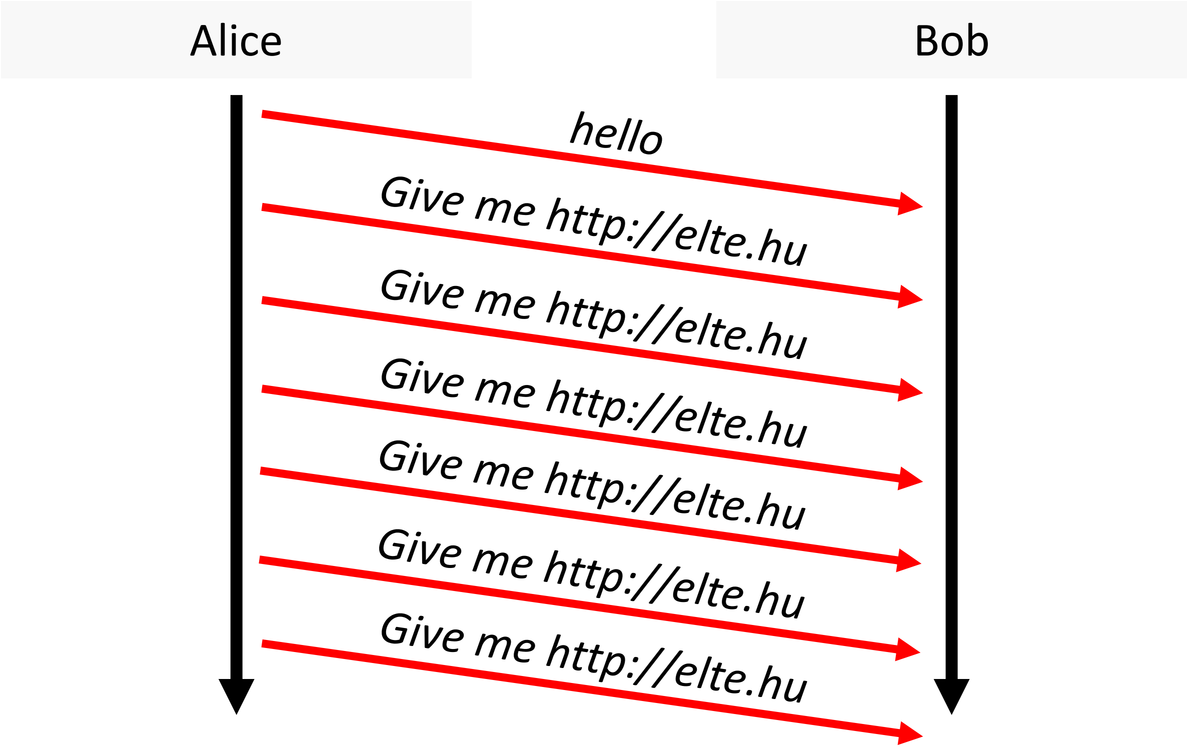
The protocol defines the order and rules the parties should follow

Who should talk next and how to respond...

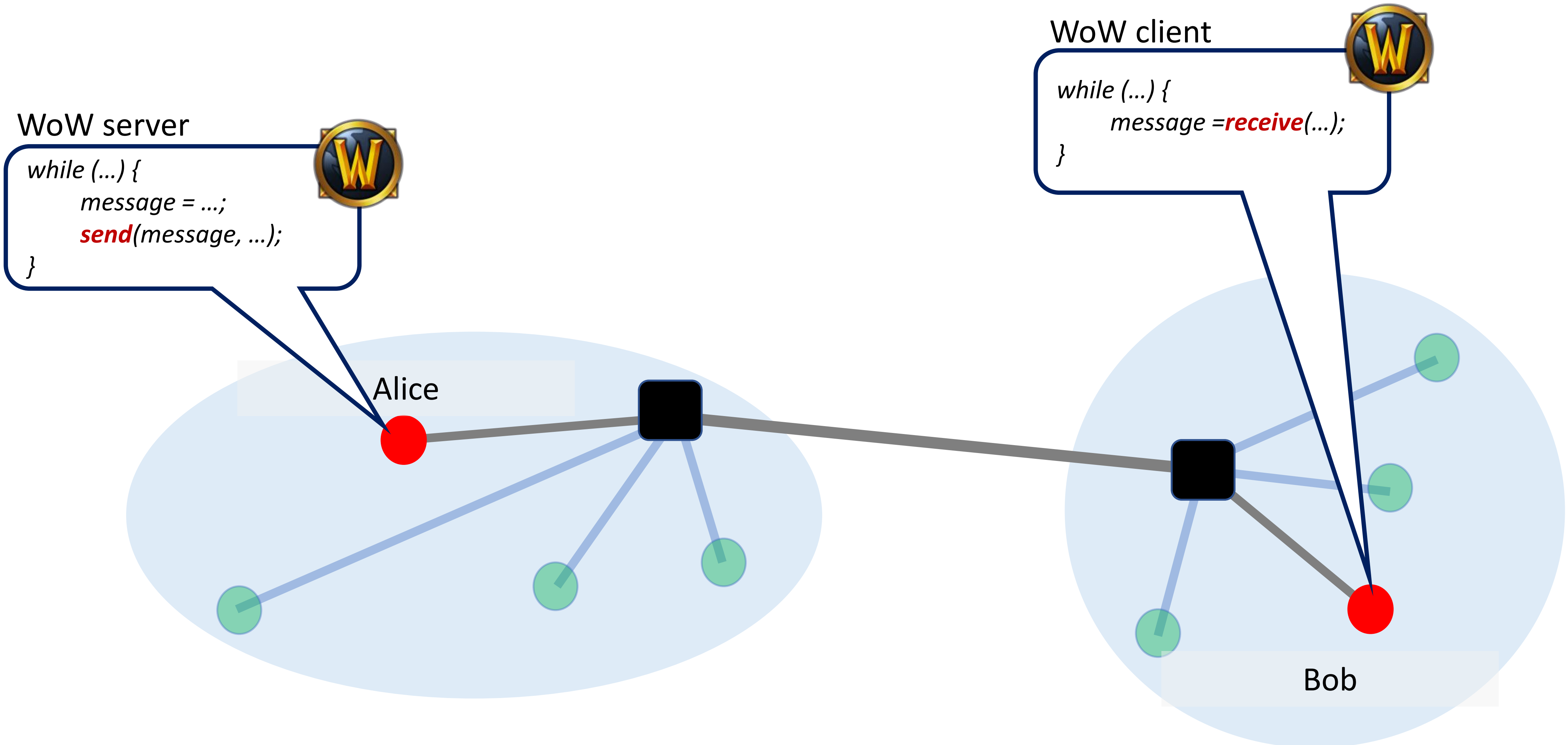


There are other kind of implementations…

Gimme, gimme, gimme
a web site after Midnight

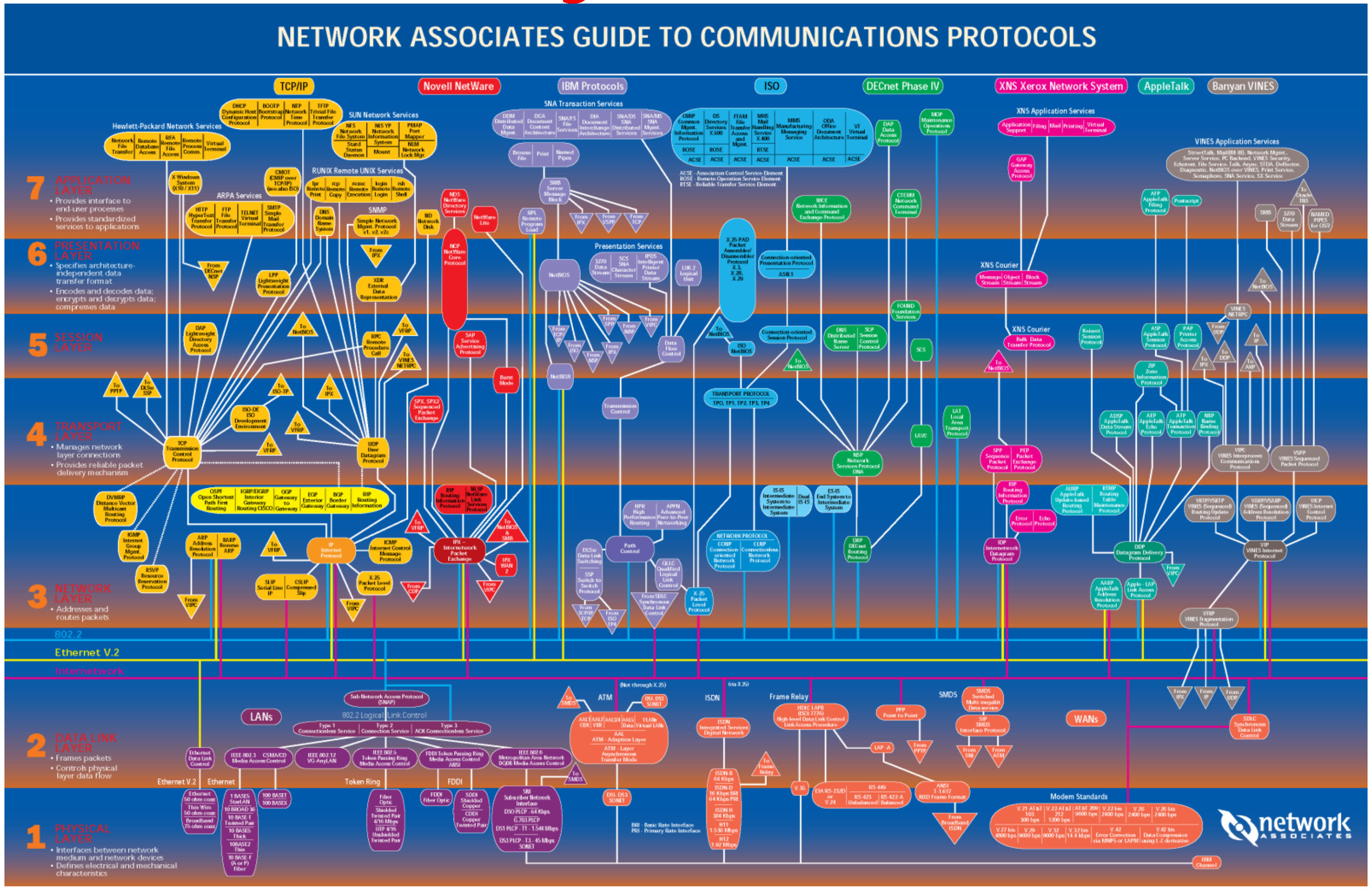


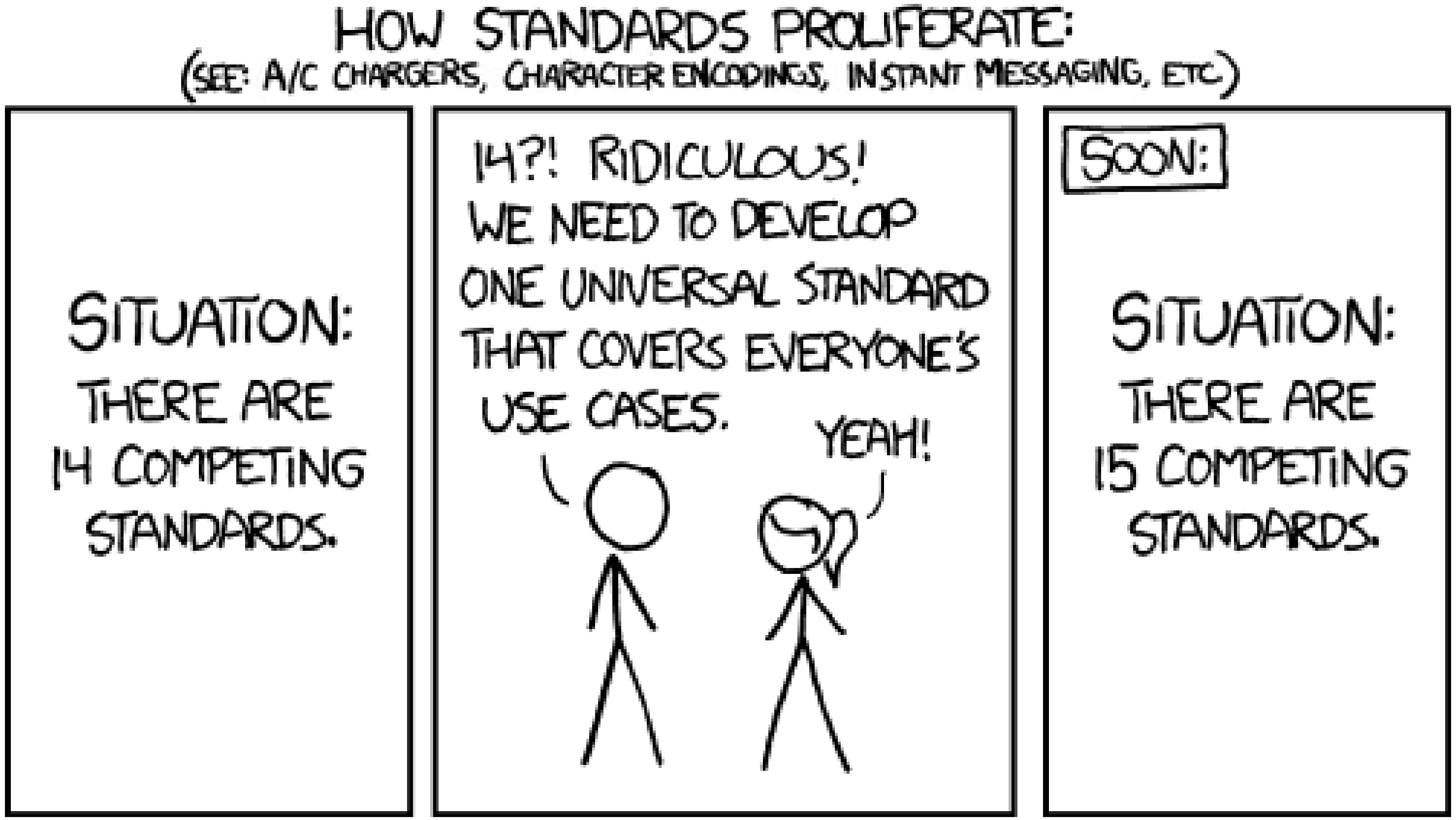
Each protocol is governed by a specific API



In practice, many existing protocols…

How does the Internet ***organize this***???




HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)
SITUATION:
THERE ARE 14 COMPETING STANDARDS.
14?! RIDICULOUS! WE NEED TO DEVELOP ONE UNIVERSAL STANDARD THAT COVERS EVERYONE'S USE CASES.
YEAH!
SOON:
SITUATION:
THERE ARE 15 COMPETING STANDARDS.

Modularity is a key component of any good system

Problem

can’t build large systems out of spaghetti code

hard (if not, impossible) to understand, debug, update

need to bound the scope of changes

evolve the system without rewriting it from scratch

Solution

Modularity is how we do it

…and understand the system at a higher-level

„Modularity,

based on abstraction,

is **the way** things get done"

Barbara Liskov, MIT

To provide structure to the design of network protocols, network designers organize protocols in layers

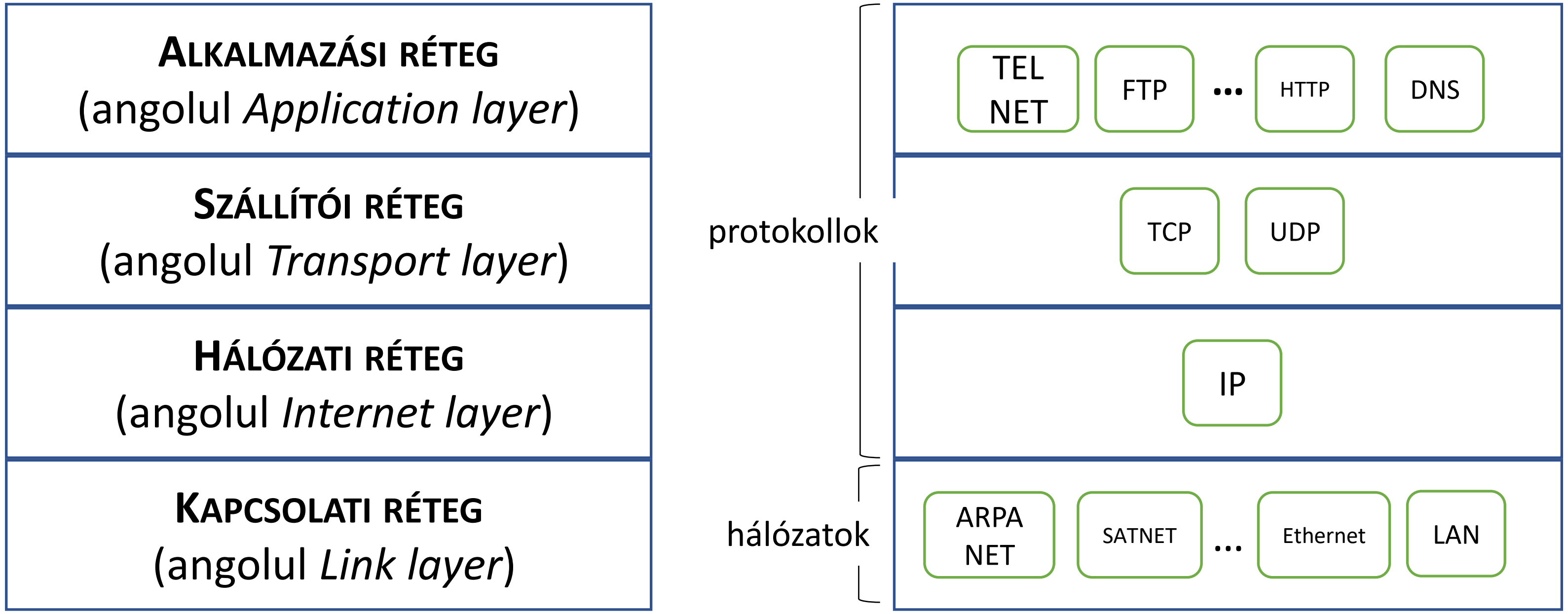
and the network hardware/software that implement them

Hálózatok modelljei

- Internet rétegmodelljei
  - TCP/IP modell: 4 réteget különböztet meg. 1982 márciusában az amerikai hadászati célú számítógépes hálózatok standardja lett. 1985-től népszerűsítették kereskedelmi felhasználásra. (*Interop*)
  - Hibrid TCP/IP modell: 5 réteget különböztet meg (*Tanenbaum, Stallings, Kurose, Forouzan*)

- Nyílt rendszerek hálózatának standard modellje
  - ***O****pen* ***S****ystem* ***I****nterconnection Reference Model*: Röviden OSI referencia modell, amely egy 7-rétegű standard, koncepcionális modellt definiál kommunikációs hálózatok belső funkcionalitásaihoz. (*ISO/IEC 7498-1*)

TCP/IP modell (RFC 1122)



TCP/IP modell rétegei („bottom-up”)

- **Kapcsolati réteg / Host-to-network or Link layer**
  - nem specifikált
  - a LAN-tól függ
- **Internet réteg / Internet or Network layer**
  - speciális csomagformátum
  - útvonal meghatározás (routing)
  - csomag továbbítás (angolul *packet forwarding*)
- **Szállítói réteg / Transport layer**
  - **T**ransport **C**ontrol **P**rotocol
    - megbízható, kétirányú bájt-folyam átviteli szolgáltatás
    - szegmentálás, folyamfelügyelet, multiplexálás
  - **U**ser **D**atagram **P**rotocol
    - nem megbízható átviteli szolgáltatás
    - nincs folyamfelügyelet
- **Alkalmazási réteg / Application layer**
  - Szolgáltatások nyújtása: Telnet, FTP, SMTP, HTTP, NNTP, DNS, SSH, etc.

Rétegek jellemzése

- Szolgáltatás
  - Mit csinál az adott réteg?
- Interfész
  - Hogyan férhetünk hozzá a réteghez?
- Protokoll
  - Hogyan implementáljuk a réteget?

Fizikai réteg

- Szolgáltatás
  - Információt visz át két fizikailag összekötött eszköz között
  - definiálja az eszköz és a fizikai átviteli közeg kapcsolatát
- Interfész
  - Specifikálja egy bit átvitelét
- Protokoll
  - Egy bit kódolásának sémája
  - Feszültség szintek
  - Jelek időzítése
- Példák: koaxiális kábel, optikai kábel, rádió frekvenciás adó

Adatkapcsolati réteg

- Szolgáltatás
  - Adatok keretekre tördelésezés: határok a csomagok között
  - Közeghozzáférés vezérlés (MAC)
  - Per-hop megbízhatóság és folyamvezérlés
- Interfész
  - Keret küldése két közös médiumra kötött eszköz között
- Protokoll
  - Fizikai címzés (pl. MAC address, IB address)
- Példák: Ethernet, Wifi, InfiniBand

Hálózati réteg

- Szolgáltatás
  - Csomagtovábbítás
  - Útvonalválasztás
  - Csomag fragmentálás kezelése
  - Csomag ütemezés
  - Puffer kezelés
- Interfész
  - Csomag küldése egy adott végpontnak
- Protokoll
  - Globálisan egyedi címeket definiálása
  - Routing táblák karbantartása
- Példák: Internet Protocol (IPv4), IPv6

Szállítói réteg

- Szolgáltatás
  - Multiplexálás/demultiplexálás
  - Torlódásvezérlés
  - Megbízható, sorrendhelyes továbbítás
- Interfész
  - Üzenet küldése egy célállomásnak
- Protokoll
  - Port szám
  - Megbízhatóság/Hiba javítás
  - Folyamfelügyelet
- Példa: UDP, TCP

Ülés v. Munkamenet réteg

- Szolgáltatás
  - kapcsolat menedzsment: felépítés, fenntarás és bontás
  - munkamenet típusának meghatározása
  - szinkronizációs pont menedzsment (checkpoint)
- Interfész
  - Attól függ…
- Protokoll
  - Token menedzsment
  - Szinkronizációs checkpoints beszúrás
- Példa: nincs

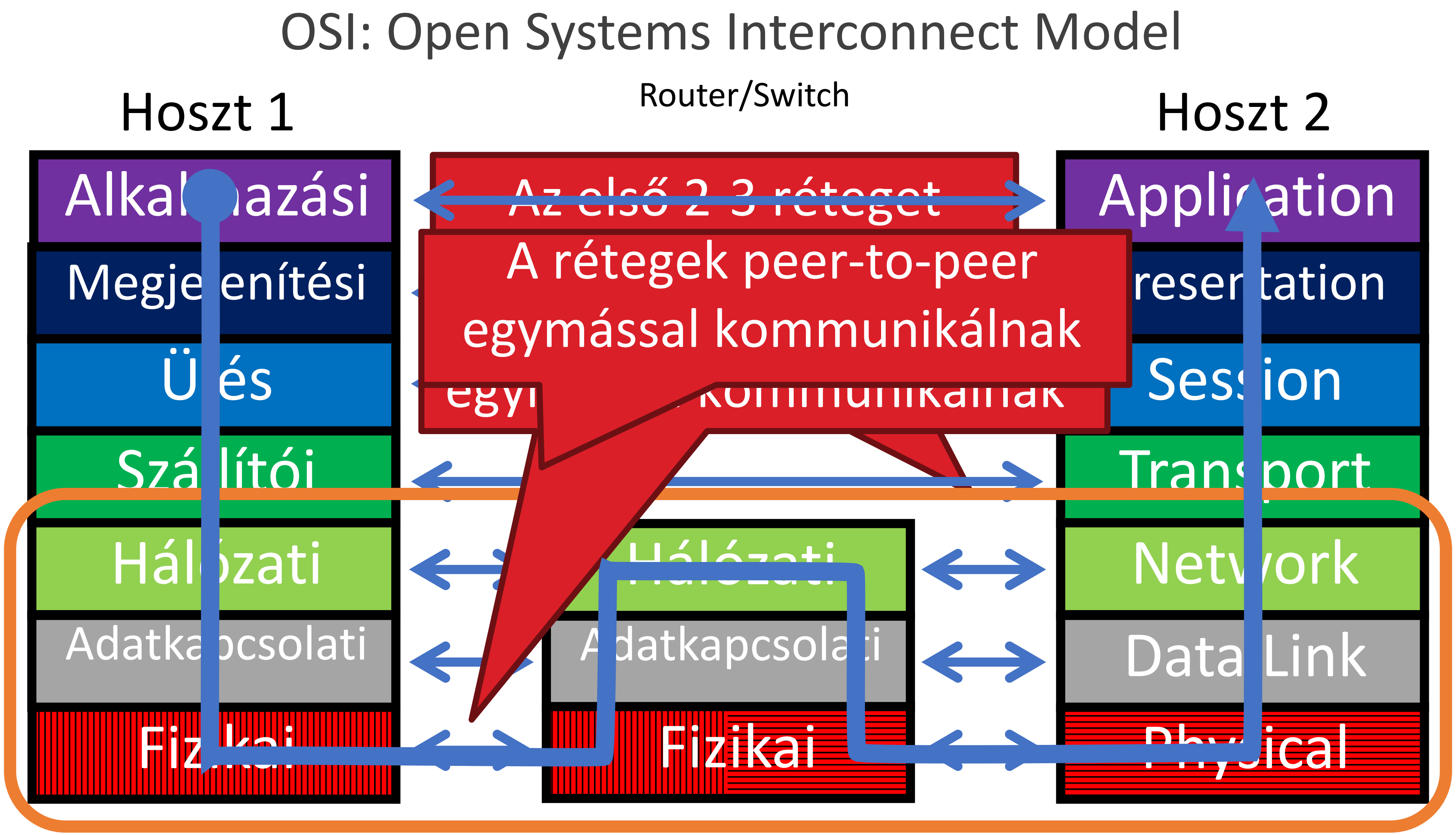
Megjelenítési réteg

- Szolgáltatás
  - Adatkonverzió különböző reprezentációk között
  - Pl. big endian to little endian
  - Pl. Ascii to Unicode
- Interfész
  - Attól függ…
- Protokoll
  - Adatformátumokat definiál
  - Transzformációs szabályokat alkalmaz
- Példa: nincs

Alkalmazási réteg

- Szolgáltatás
  - Bármi…
- Interfész
  - Bármi…
- Protokoll
  - Bármi…
- Példa: kapcsold be a mobilod és nézd meg milyen appok vannak rajta…

ISO OSI modell



Hybrid model – 5 layers

Each layer provides a service to the layer above

	layer	service provided
L5	**Application**	*high level network access*
L4	**Transport**	*end-to-end delivery (reliable or not)*
L3	**Network**	*global best-effort delivery*
L2	**Link**	*local best-effort delivery*
L1	**Physical**	*physical transfer of bits*

Each layer provides a service to the layer above by using the services of the layer directly below it

Applications

...built on...

Reliable (or unreliable) transport

...built on...

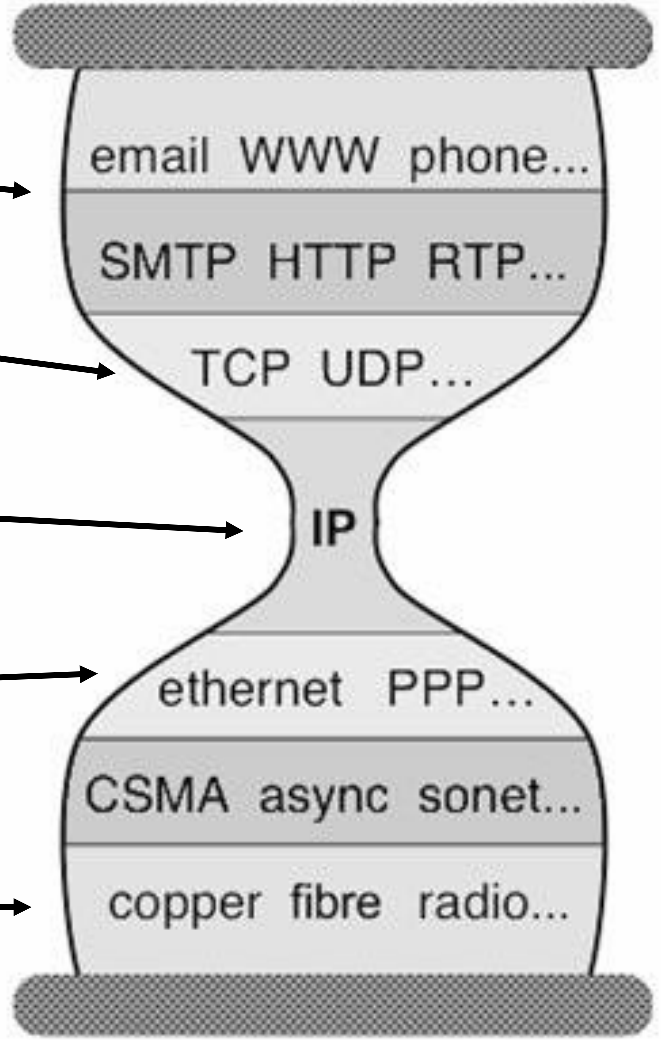
Best-effort global packet delivery

...built on...

Best-effort local packet delivery

...built on...

Physical transfer of bits



Each layer has a unit of data (aka protocol data unit)

	layer	role (PDU)
L5	**Application**	exchanges messages between processes
L4	**Transport**	transports segments between end-systems
L3	**Network**	moves packets around the network
L2	**Link**	moves frames across a link
L1	**Physical**	moves bits across a physical medium

Each layer (except for L3) is implemented with different protocols

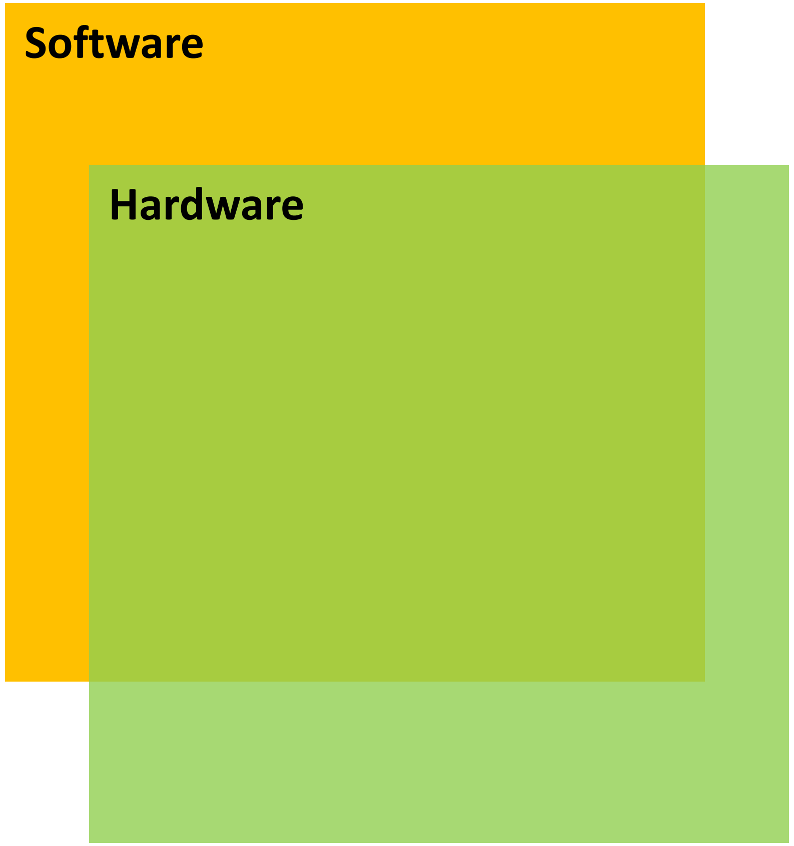
	layer	protocols
L5	**Application**	HTTP, SMTP, FTP, SIP, …
L4	**Transport**	TCP, UDP, SCTP
L3	**Network**	IP
L2	**Link**	Ethernet, Wifi, ADSL, WiMAX, LTE, …
L1	**Physical**	Twisted pair, fiber, coaxial cable, …

The Internet Protocol (IP) is the glue acting as a unifying network layer

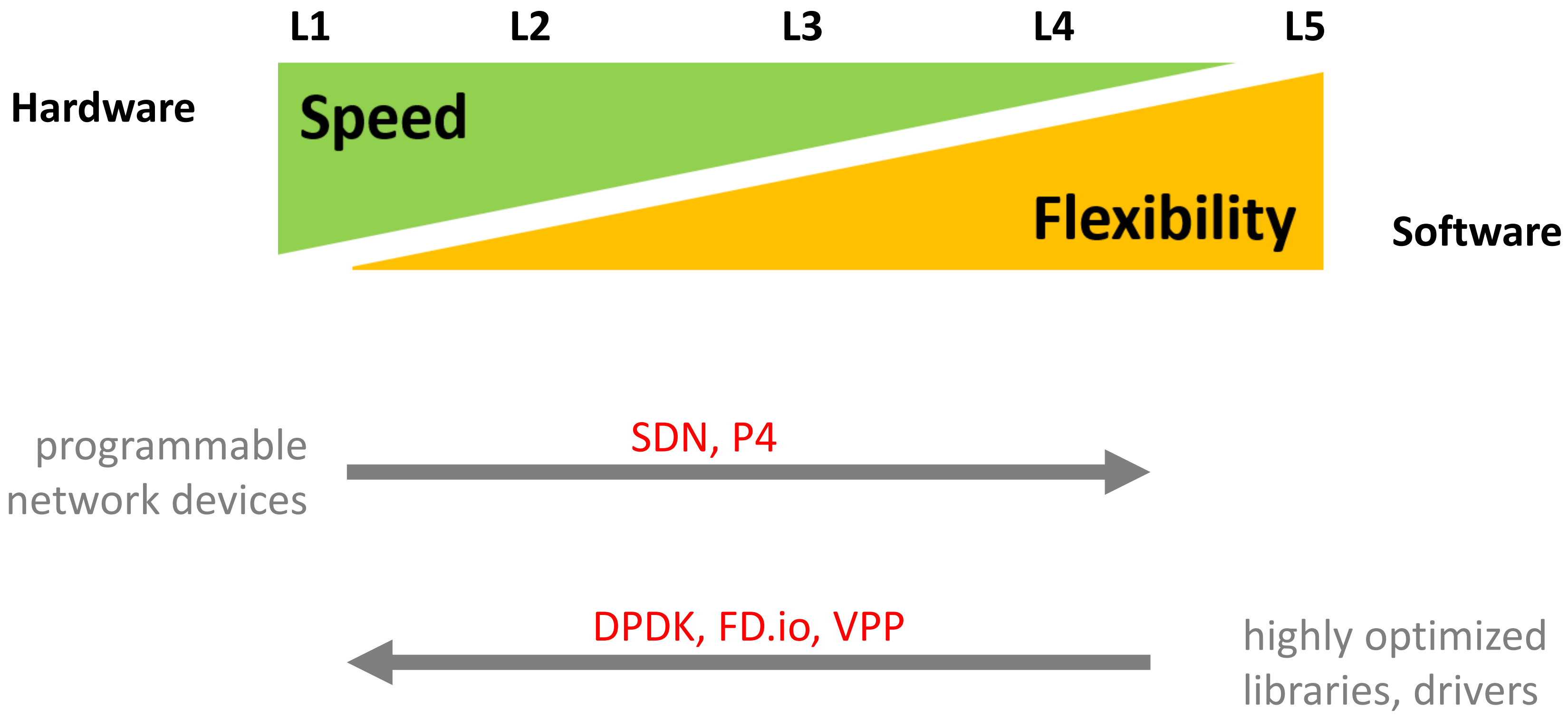
	layer	protocols
L5	**Application**	HTTP, SMTP, FTP, SIP, …
L4	**Transport**	TCP, UDP, SCTP
L3	**Network**	IP
L2	**Link**	Ethernet, Wifi, ADSL, WiMAX, LTE, …
L1	**Physical**	Twisted pair, fiber, coaxial cable, …

Each layer is implemented with different protocols and technologies

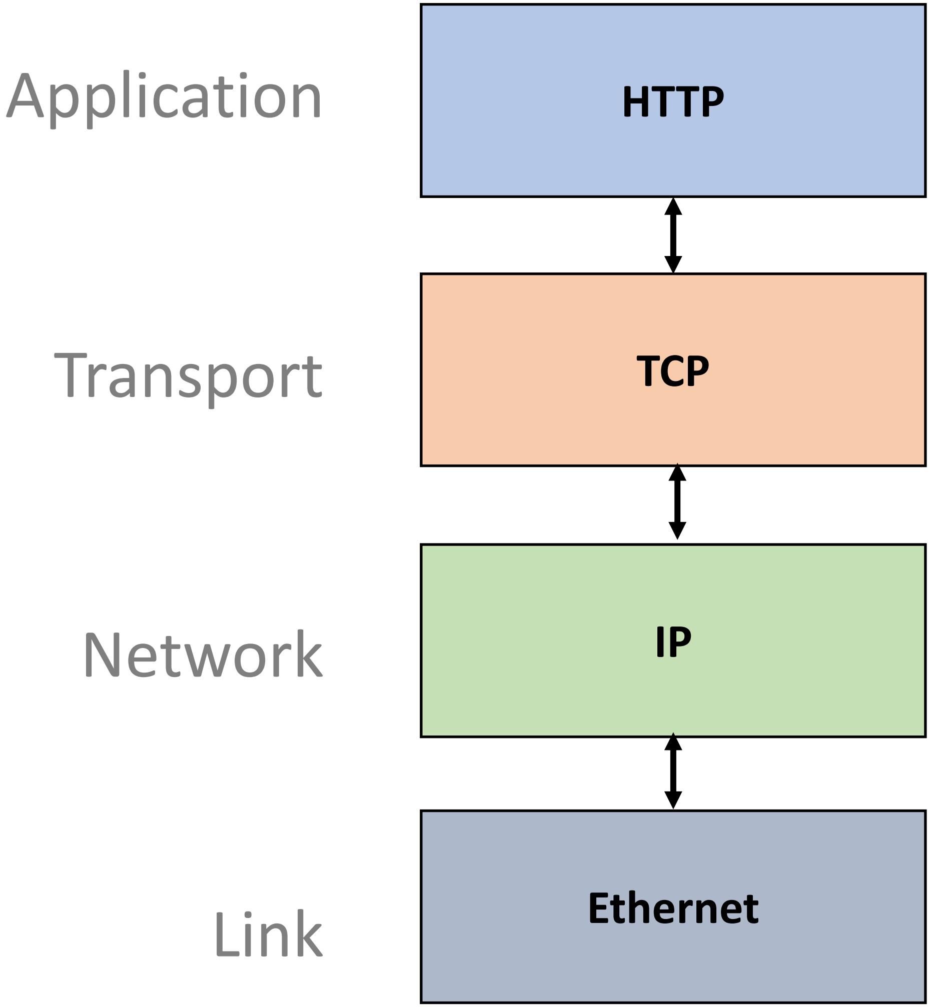
	layer	technology
L5	**Application**	
L4	**Transport**	
L3	**Network**	
L2	**Link**	
L1	**Physical**	

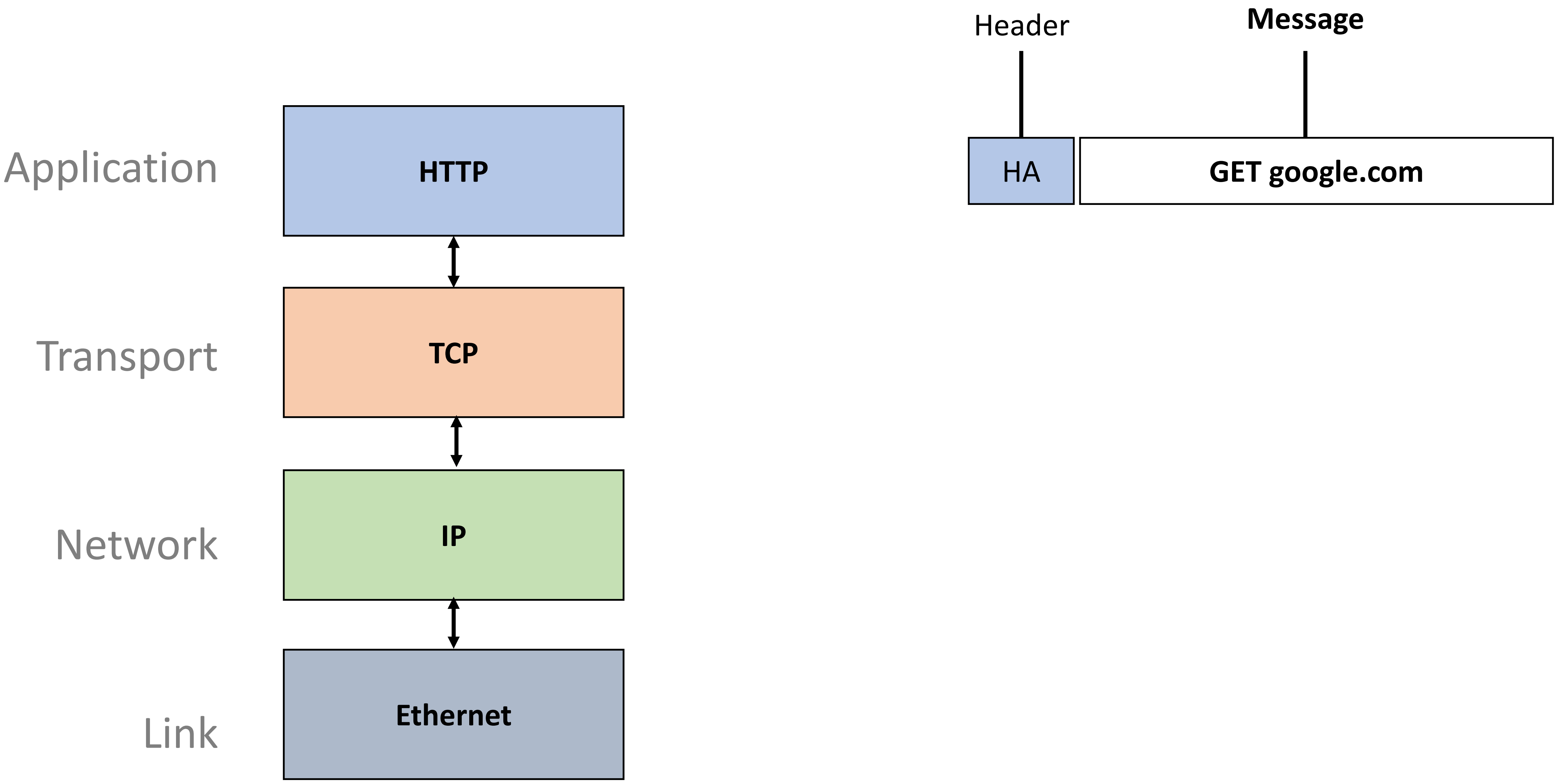


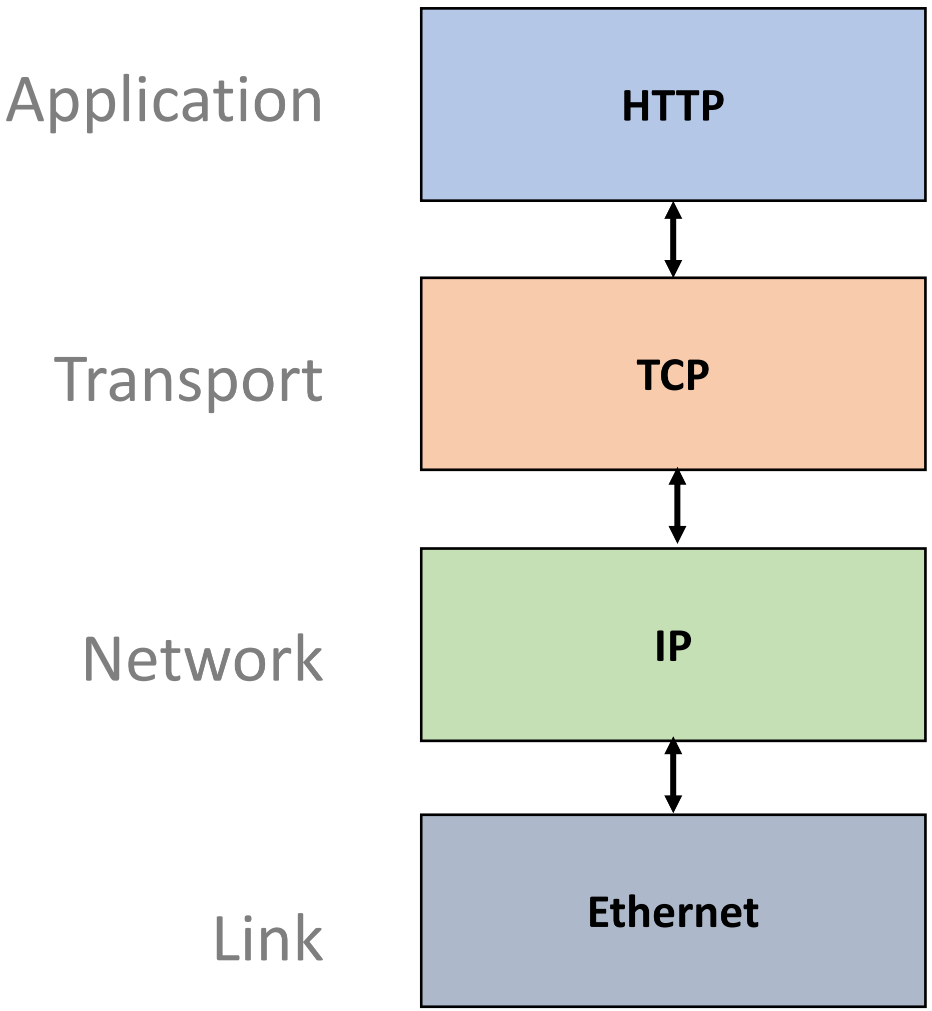
Software and hardware advancements

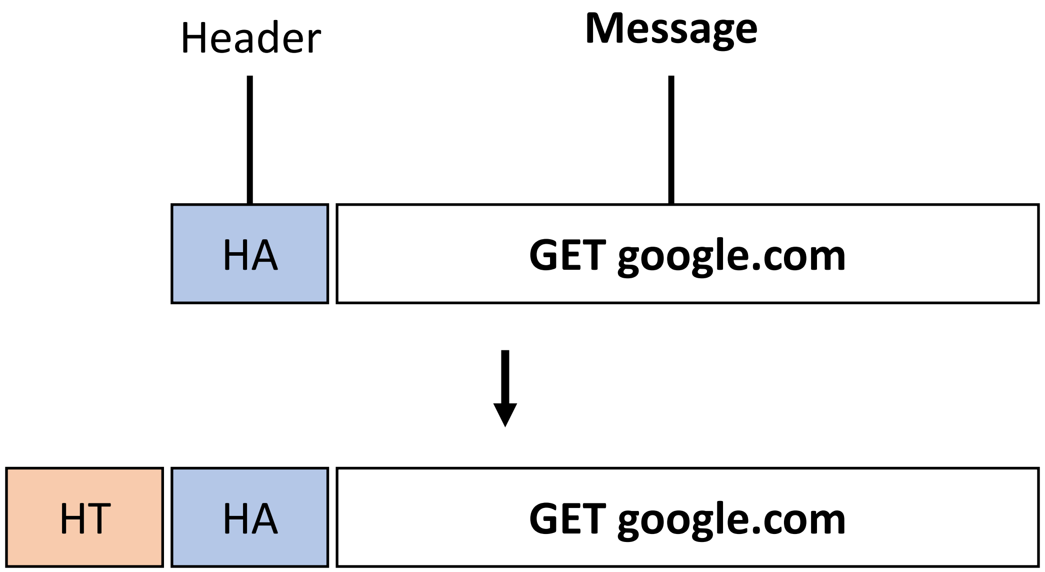


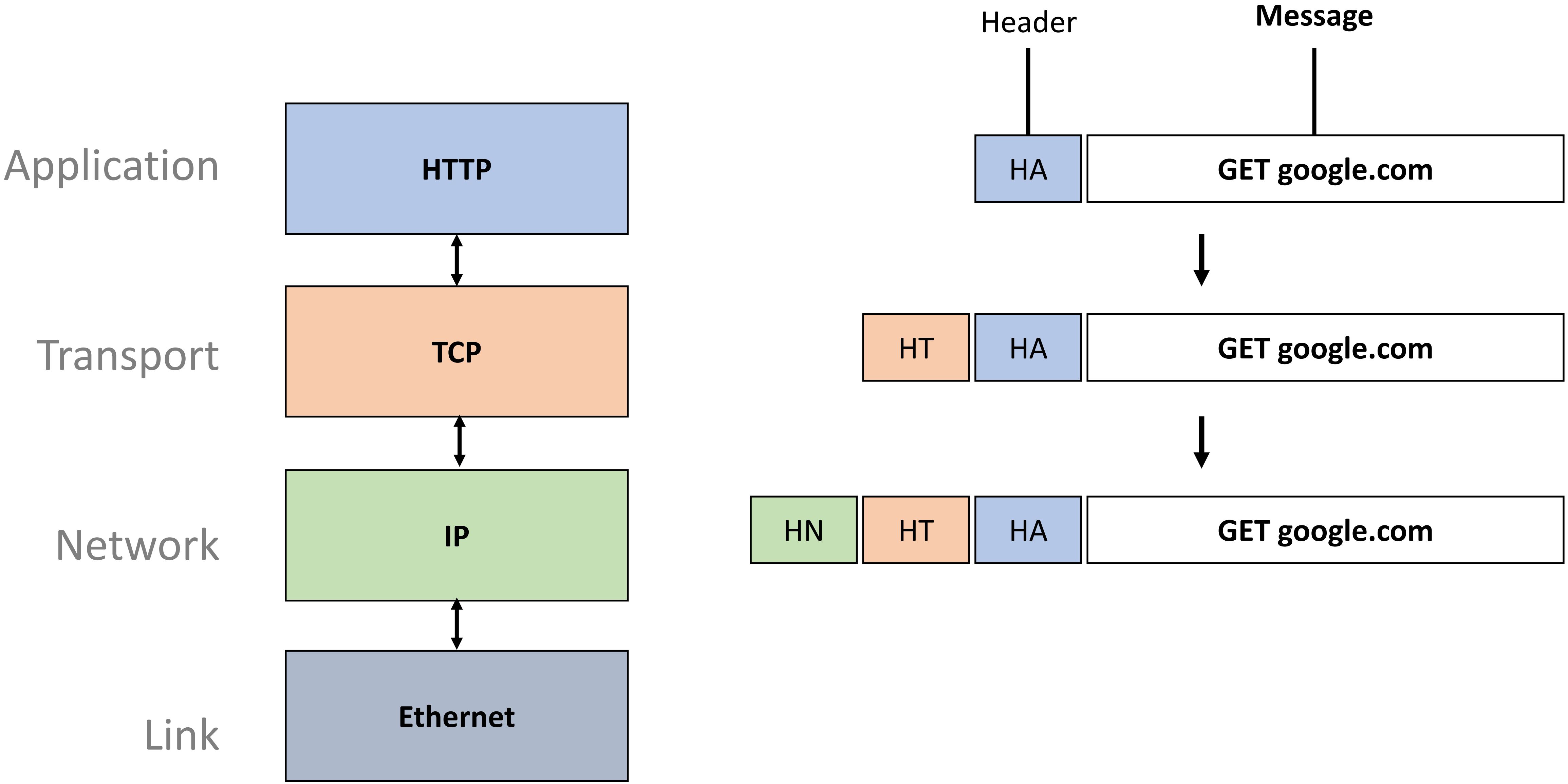
Each layer takes messages from the layer above, and *encapsulates* with its own header and/or trailer



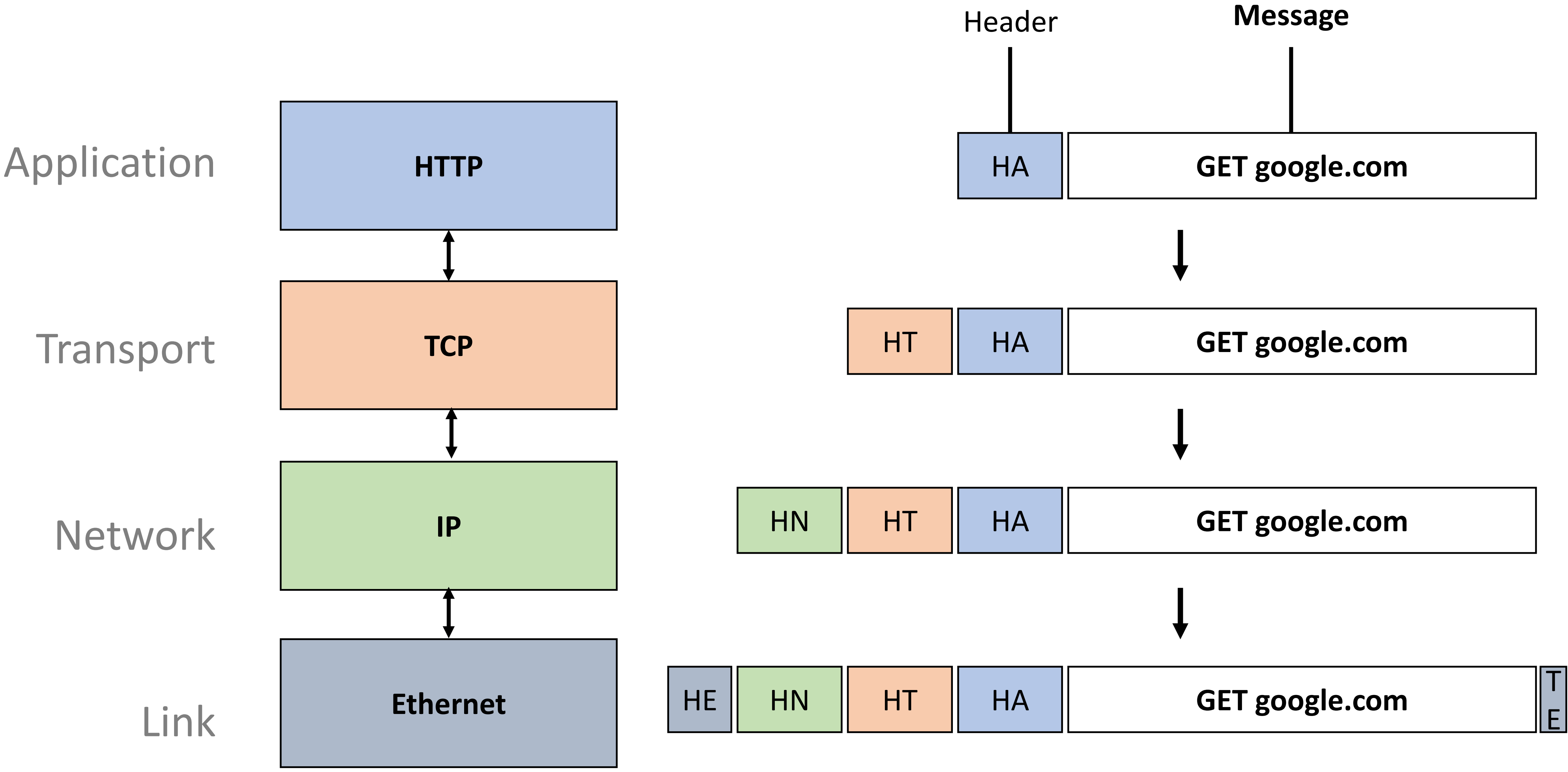

Application
HTTP
Transport
TCP
Network
IP
Link
Ethernet
Header
Message
HA
GET google.com


Application
HTTP
Transport
TCP
Network
IP
Link
Ethernet


Header
Message
HA
GET google.com
HT
HA
GET google.com

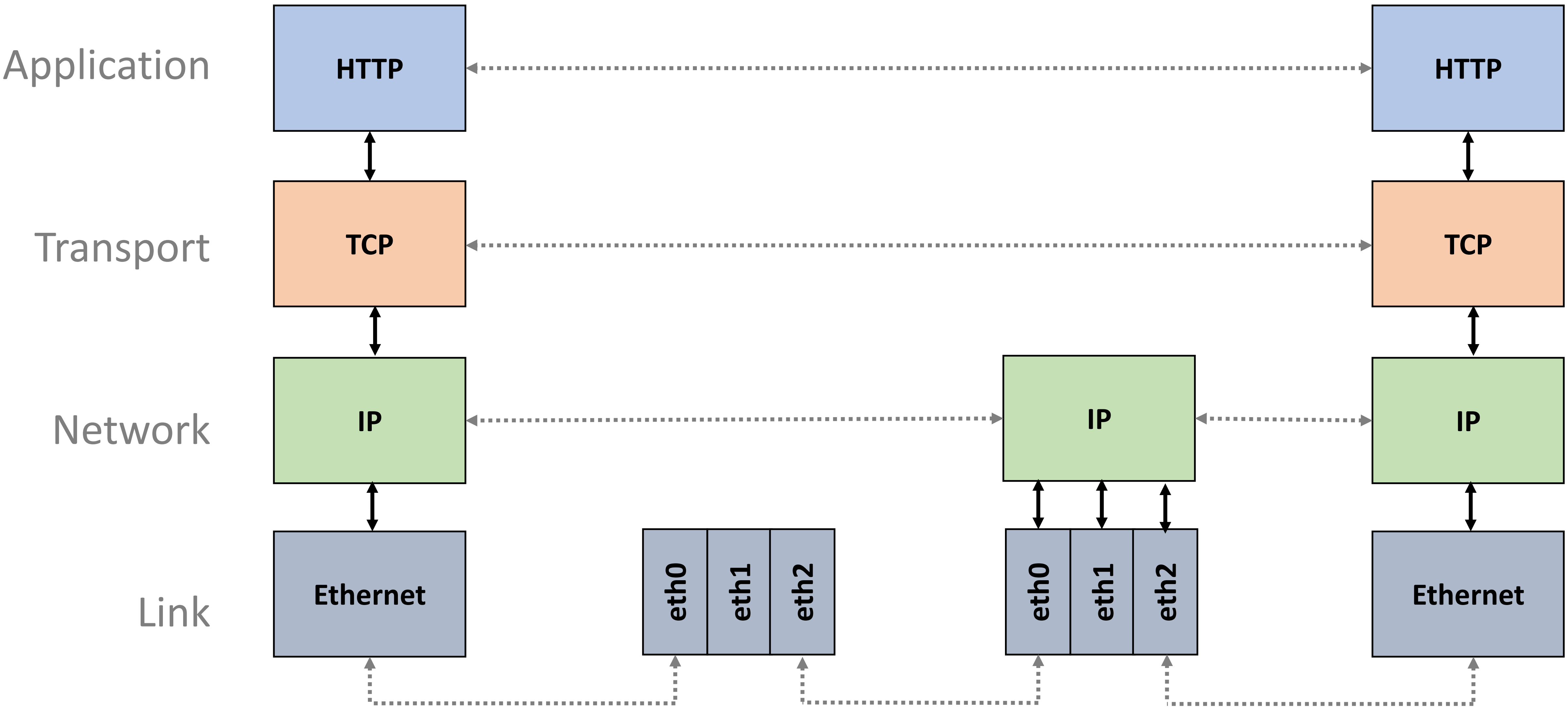


Application
HTTP
Transport
TCP
Network
IP
Link
Ethernet
Header
Message
HA
GET google.com
HT
HA
GET google.com
HN
HT
HA
GET google.com

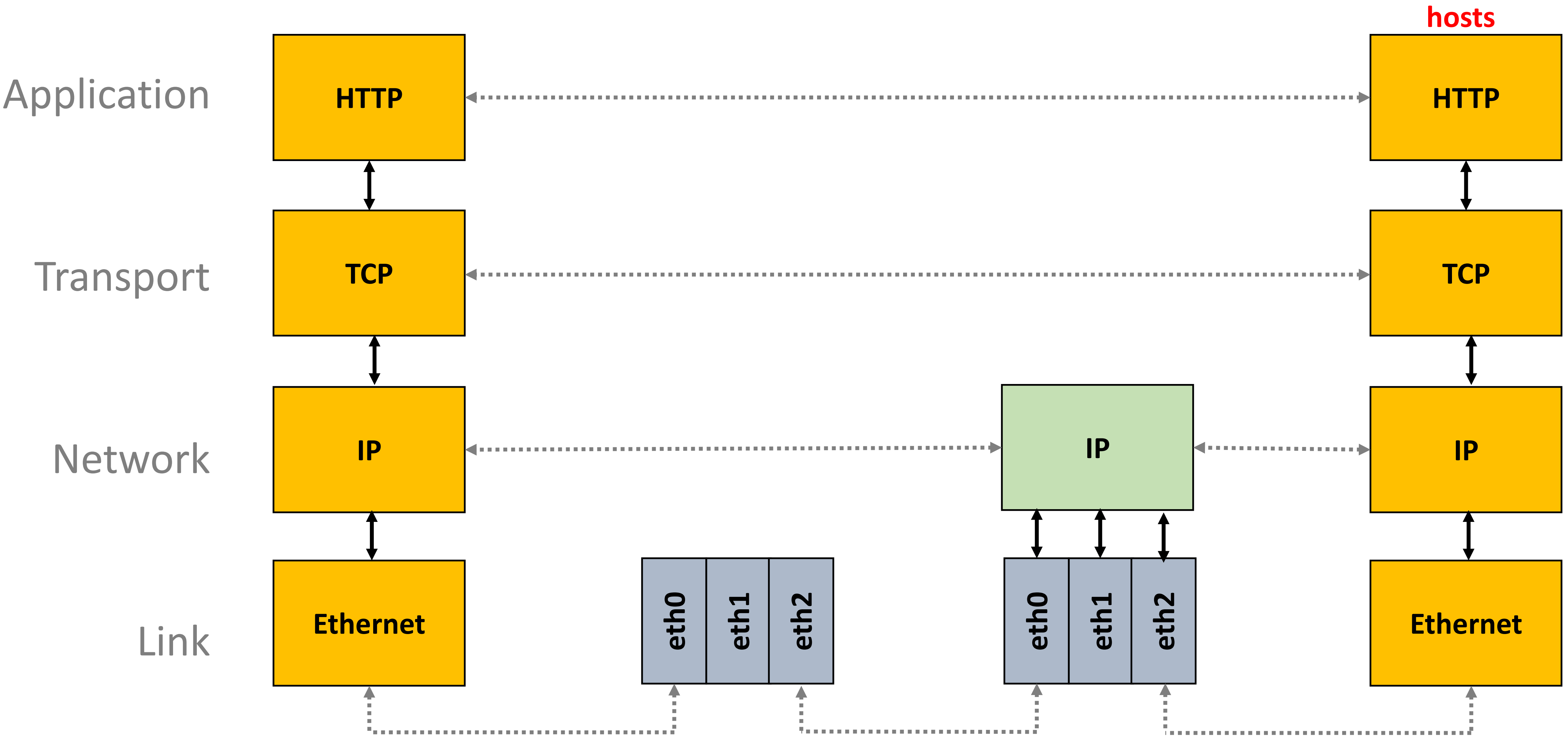


Header
Message
Application
HTTP
HA
GET google.com
Transport
TCP
HT
HA
GET google.com
Network
IP
HN
HT
HA
GET google.com
Link
Ethernet
HE
HN
HT
HA
GET google.com
TE

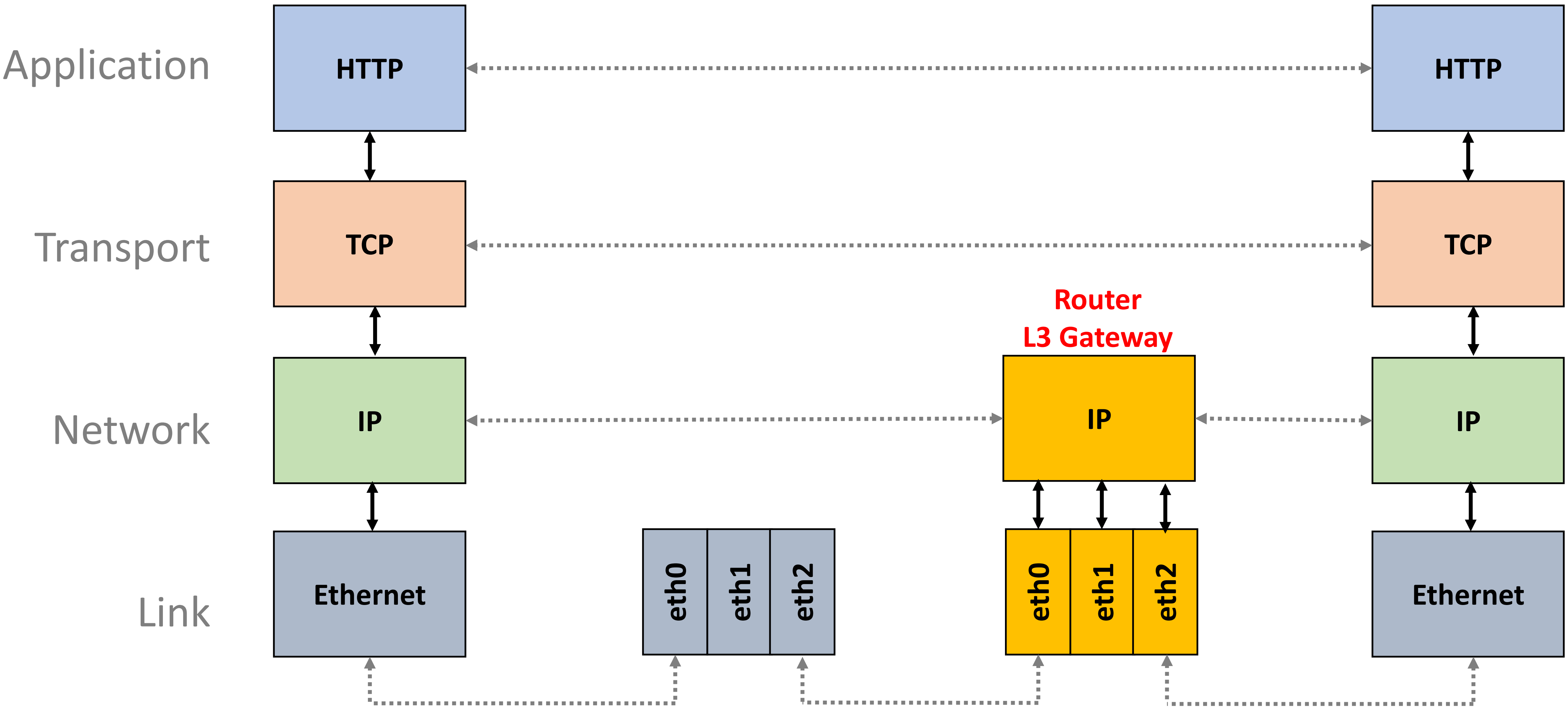
In practice, layers are **distributed** on every network device



Since when bits arrive they must make it to the application, all the layers exist on a host

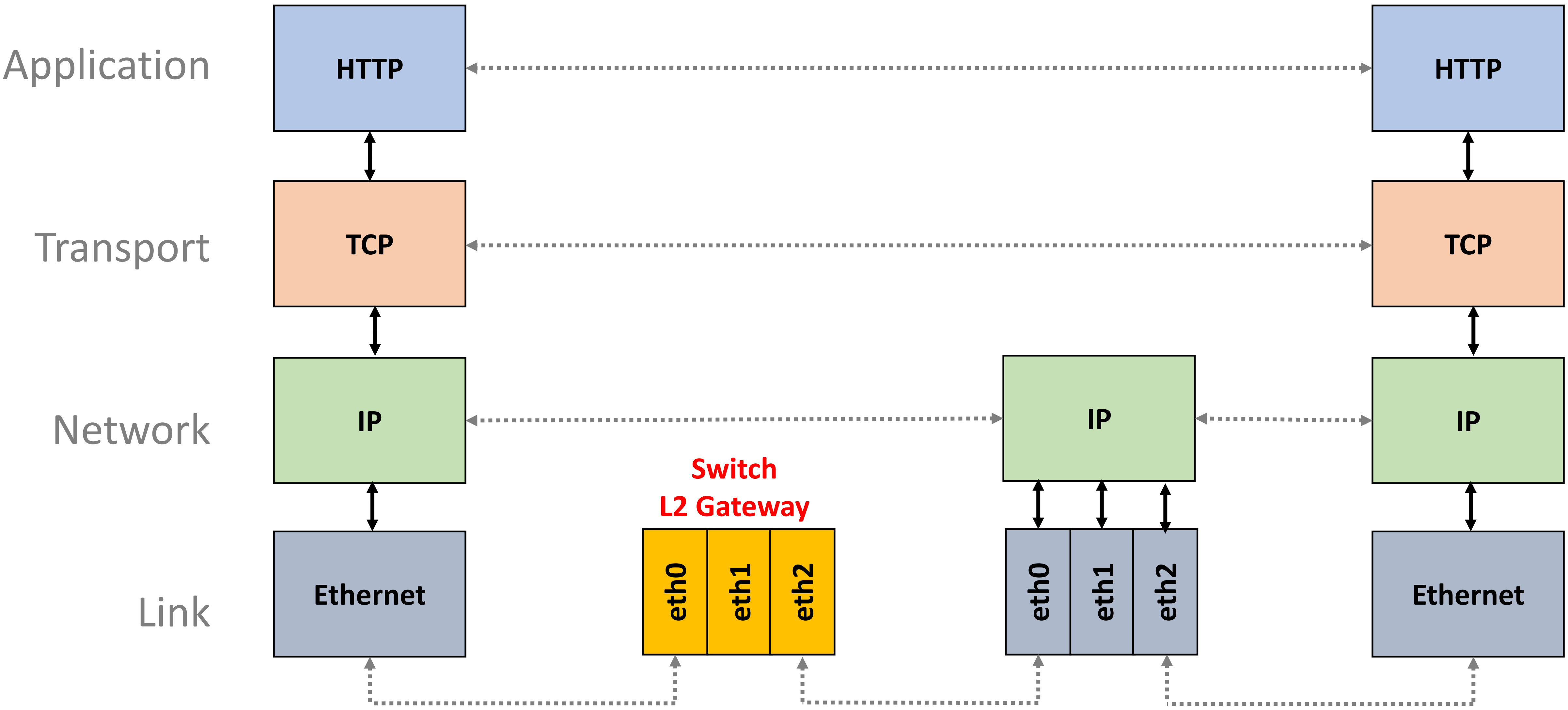


Routers act as L3 gateway
as such they implement L2 and L3



Switches act as L2 gateway

as such they only implement L2



Overview

How do we characterize the network?

A network *connection* is characterized by its delay, loss rate and throughput

delay

loss

throughput

How long does it take for a packet to reach the destination

What fraction of packets sent to a destination are dropped?

At what rate is the destination receiving data from the source?

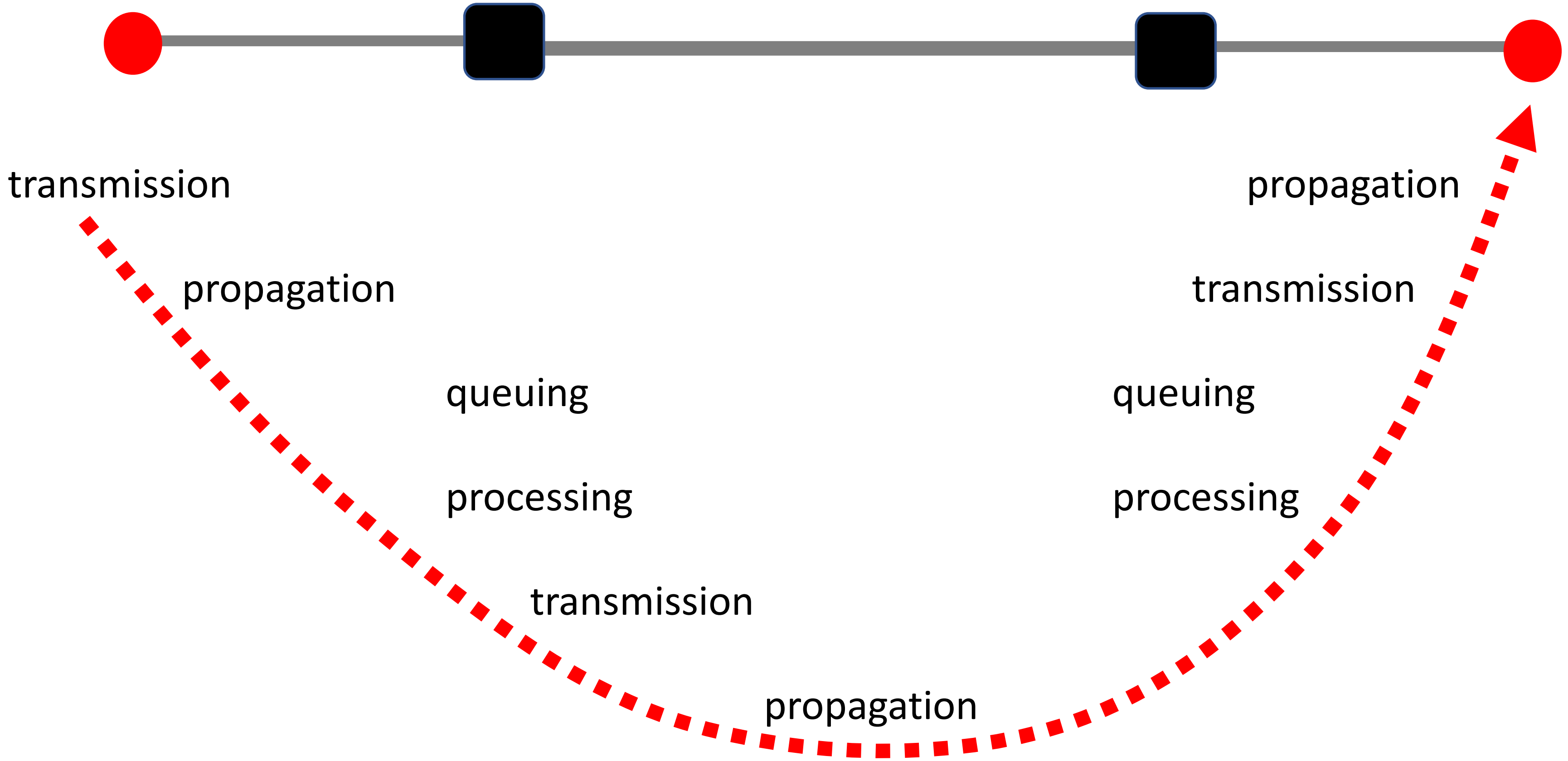
Delay

Each packet suffers from several types of delays at *each node* along the path

	transmission delay	due to **link properties**
+	**propagation** delay	
+	**processing** delay	due to **traffic mix & switch internals**
+	**queueing** delay	
=	**total** delay	

Each packet suffers from several types of delays at *each node* along the path

	transmission delay
+	**propagation** delay
~~+~~	~~**processing** delay~~ tend to be tiny
+	**queueing** delay
=	**total** delay


transmission
propagation
queuing
processing
transmission
propagation
queuing
processing
transmission
propagation

The transmission delay is the amount of time required to push all of the bits onto the link

$$\text{Transmission delay [sec]} = \frac{\text{packet size [\#bits]}}{\text{link bandwidth [\#bits/sec]}}$$

$$\text{Example} = \frac{1000 \text{ bits}}{100 \text{ Mpbs}} = 10\ \mu\text{sec}$$

The propagation delay is the amount of time required for a bit to travel to the end of the link

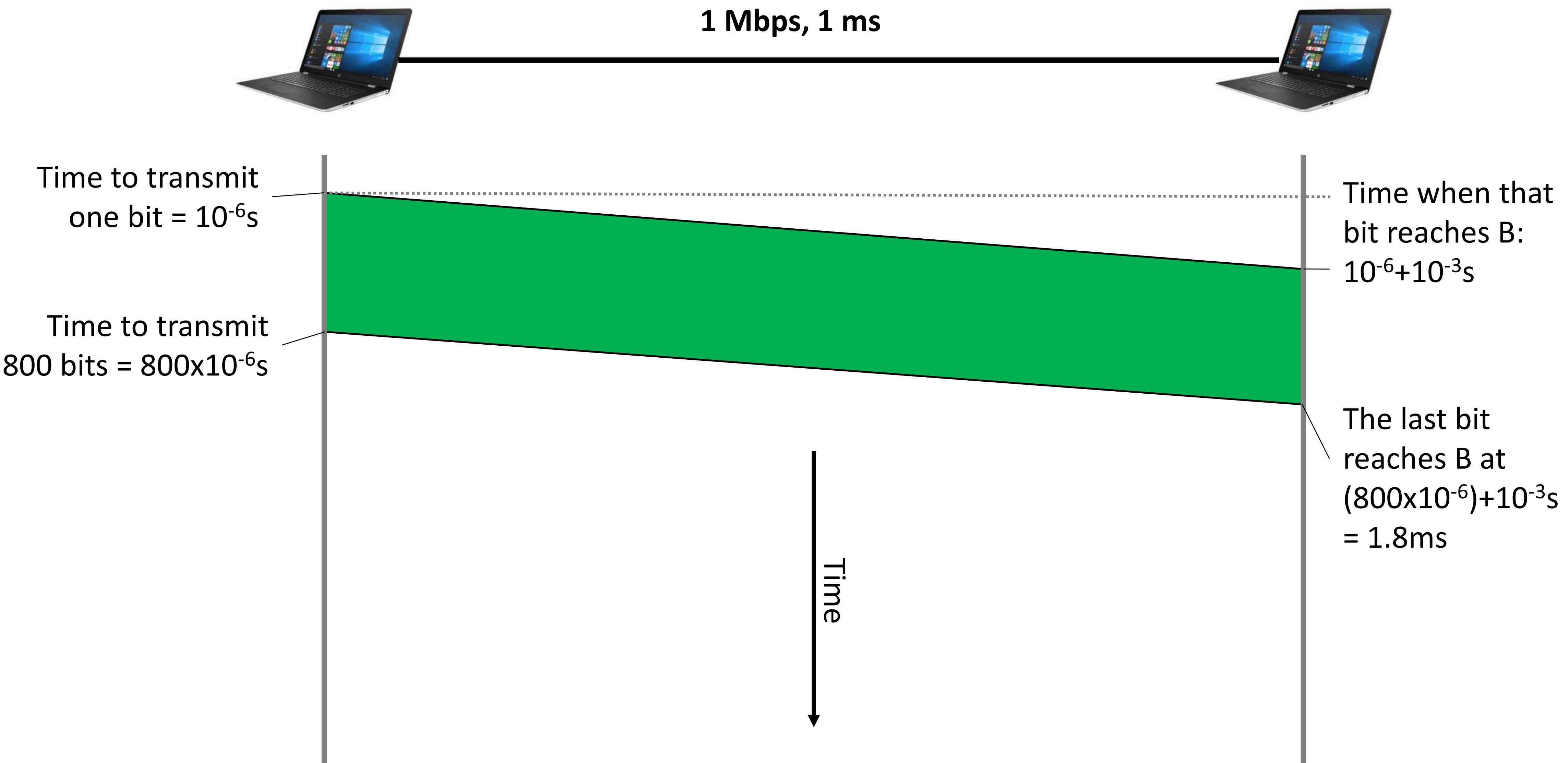
$$\underset{[\text{sec}]}{\text{Propagation delay}} = \frac{\text{link length} \quad [\text{m}]}{\text{signal propagation speed} \quad [\text{m/sec}]}$$

Example

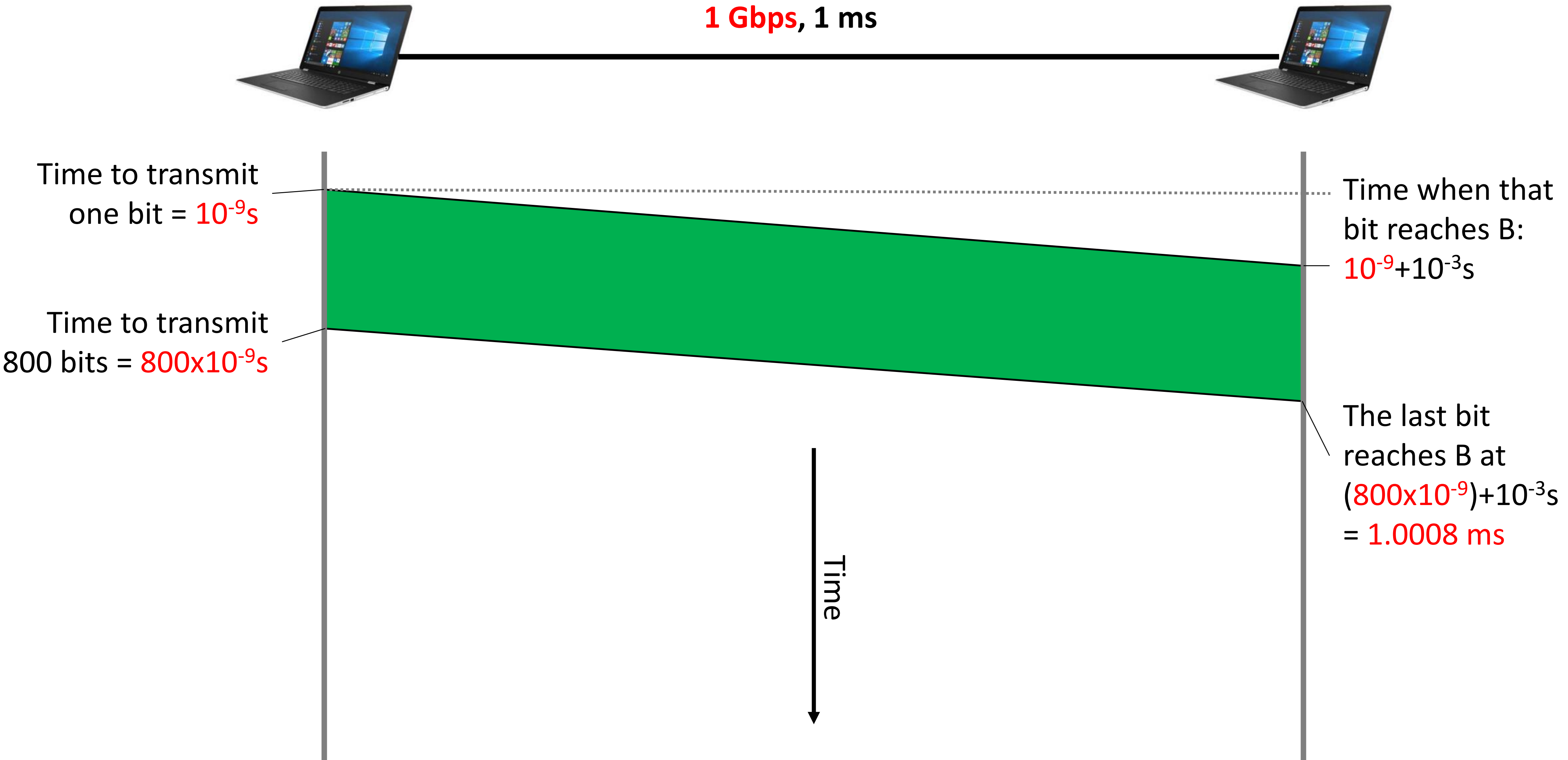
$$= \frac{30000 \text{ m}}{2\text{x}10^{8} \text{ m/sec}} = 150\ \mu\text{sec}$$

(speed of light in fiber)

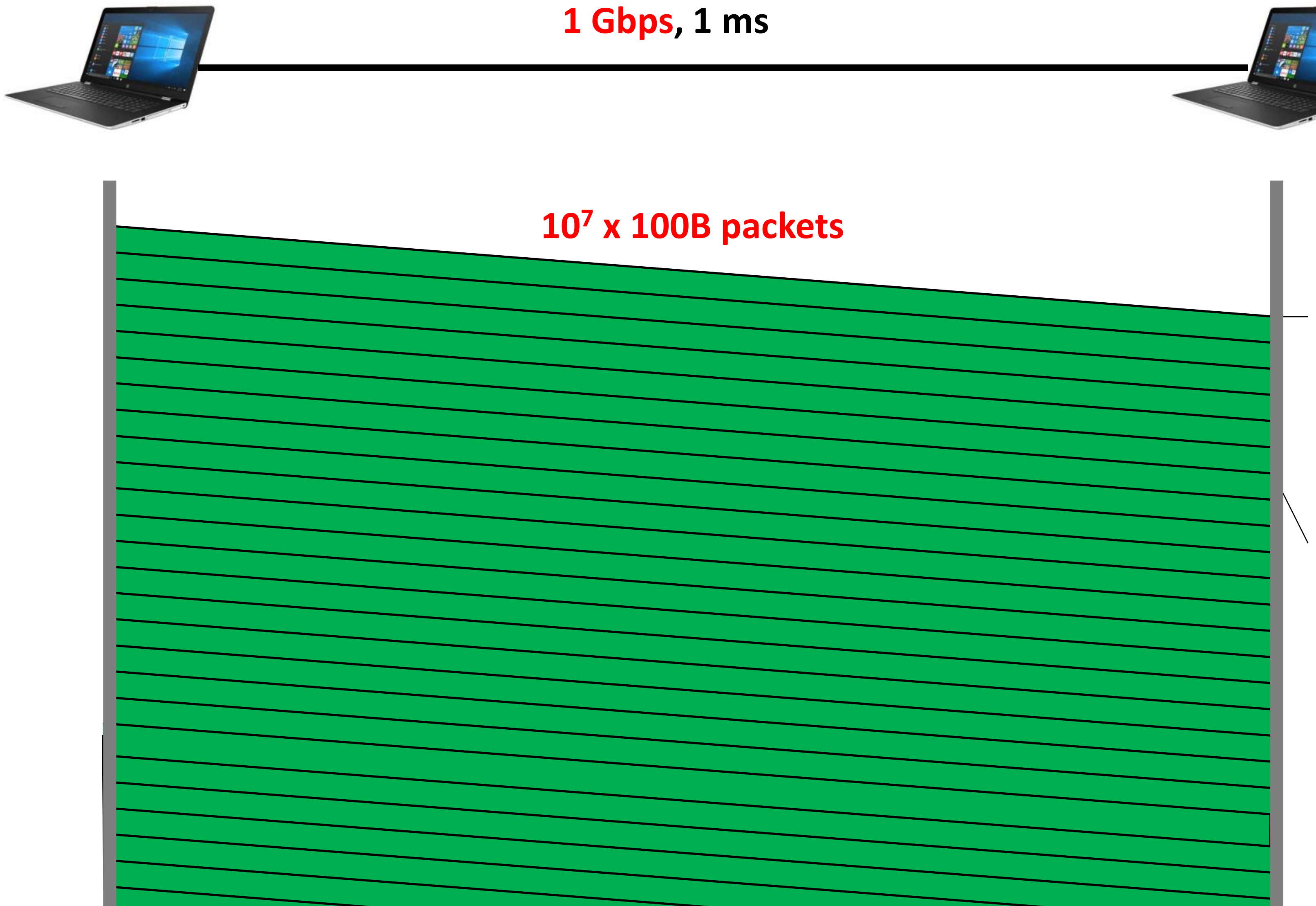
How long does it take to exchange 100 Bytes packet?



If we have a 1 Gbps link, the total time decreases to 1.0008 ms



If we now exchange a 1GB file split in 100B packets



Time when that bit reaches B: $10^{-9}+10^{-3}$s

The last bit reaches B at $(10^7 \text{x} 800 \text{x} 10^{-9})+10^{-3}$s = 8001 ms

Different transmission characteristics imply different tradeoffs in terms of which delay dominates

10^7x100B pkts	1Gbps link	***transmission delay dominates***
1x100B pkt	1Gbps link	***propagation delay dominates***
1x100B pkt	1Mbps link	***both matter***

In the Internet, we cannot know in advance which one matter!

The queuing delay is the amount of time a packet waits (in a buffer) to be transmitted on a link

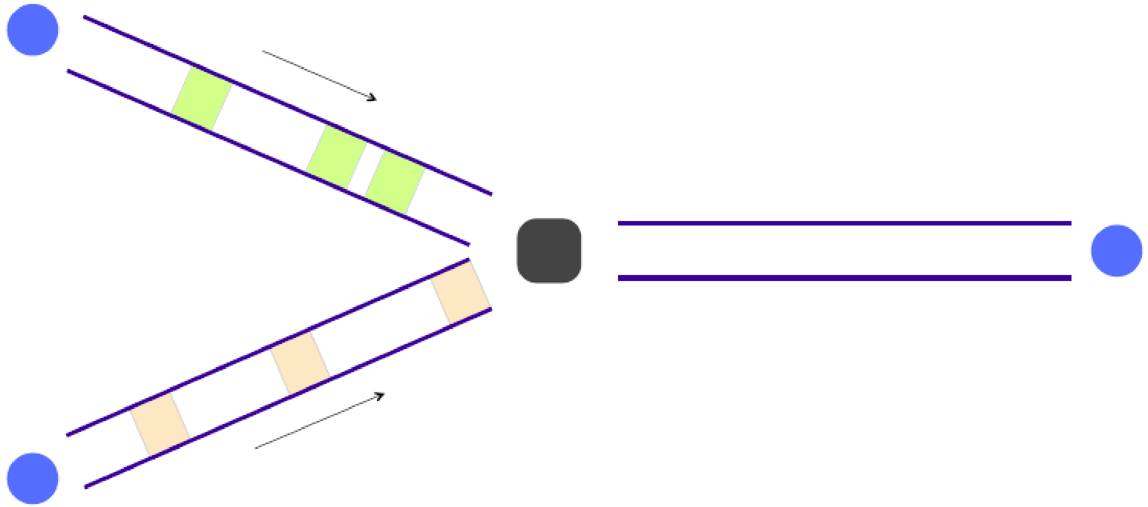
Queuing delay is the hardest to evaluate

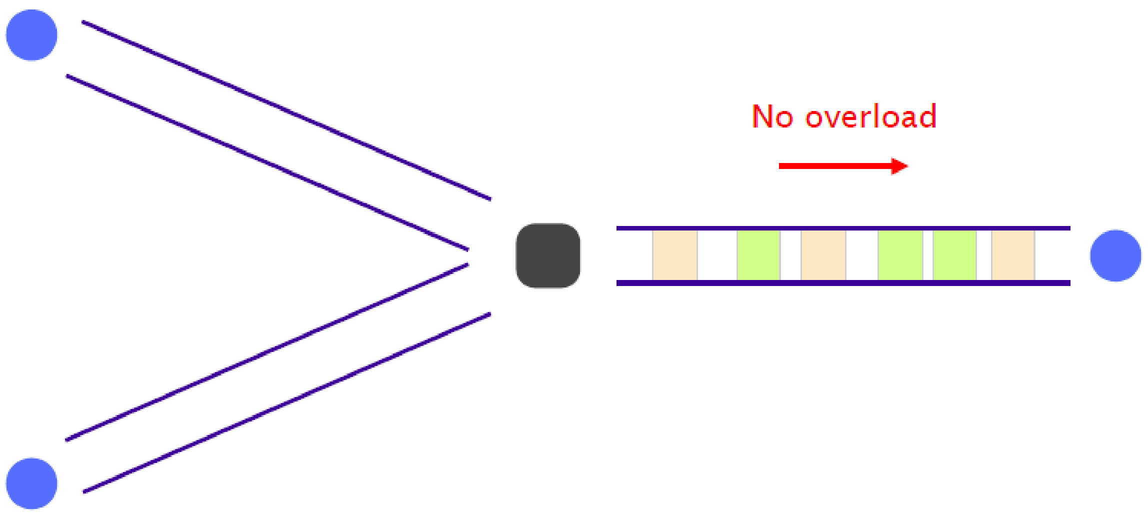
as it varies from packet to packet

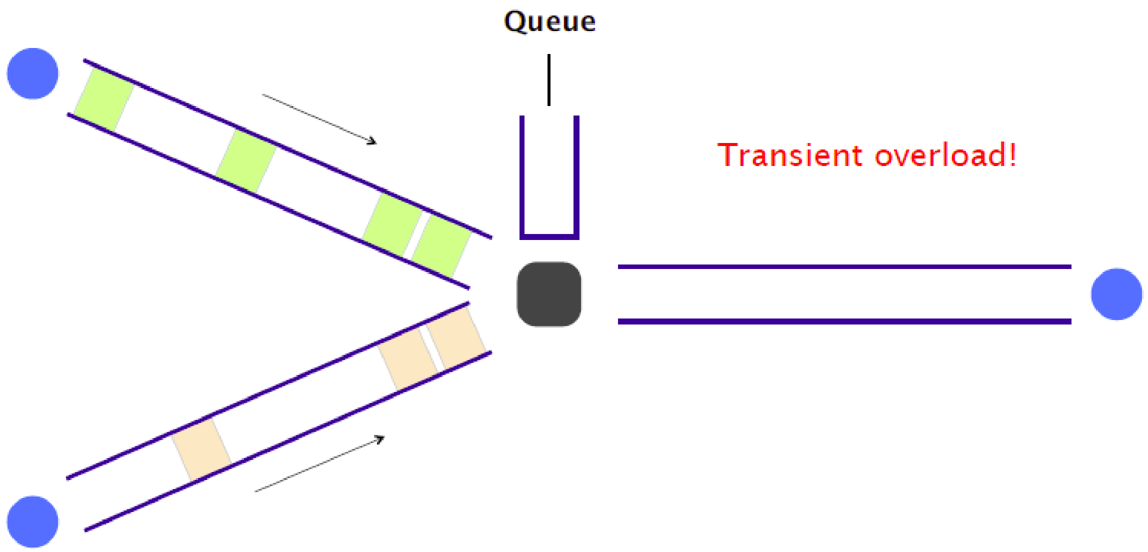
It is characterized with statistical measures

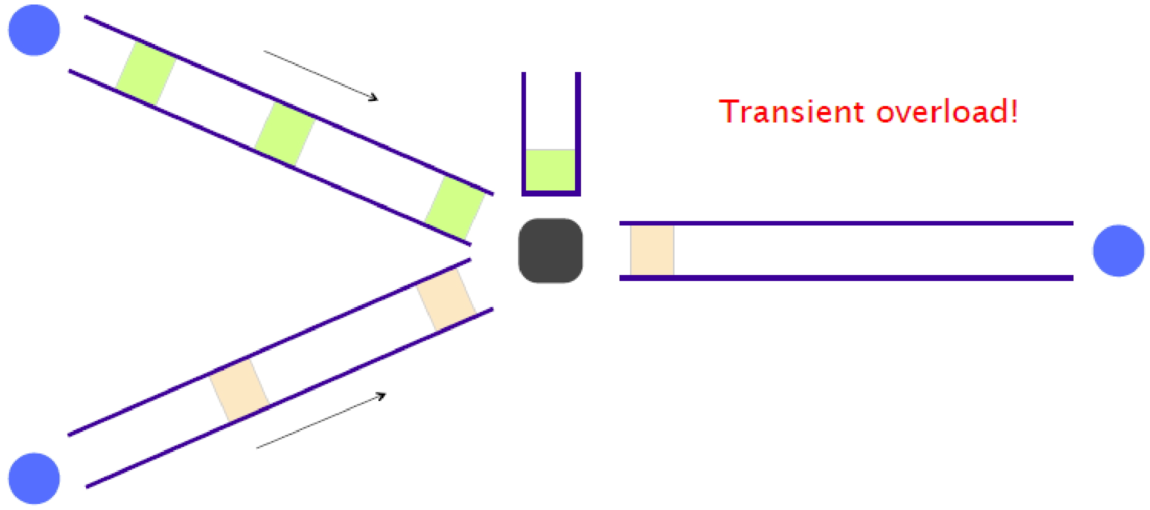
e.g., average delay & variance, probability of exceeding x

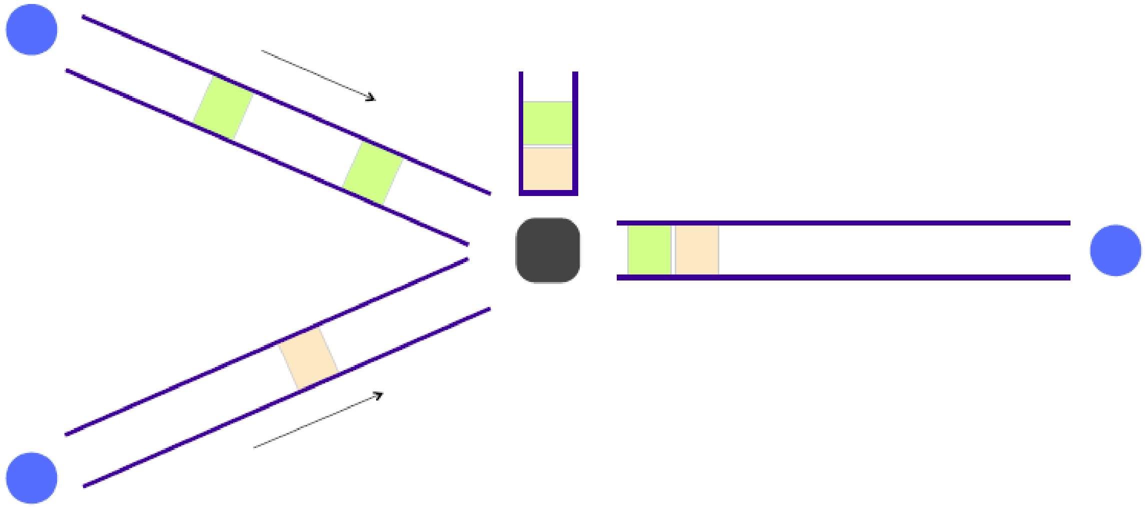
Queuing delay depends on the traffic pattern

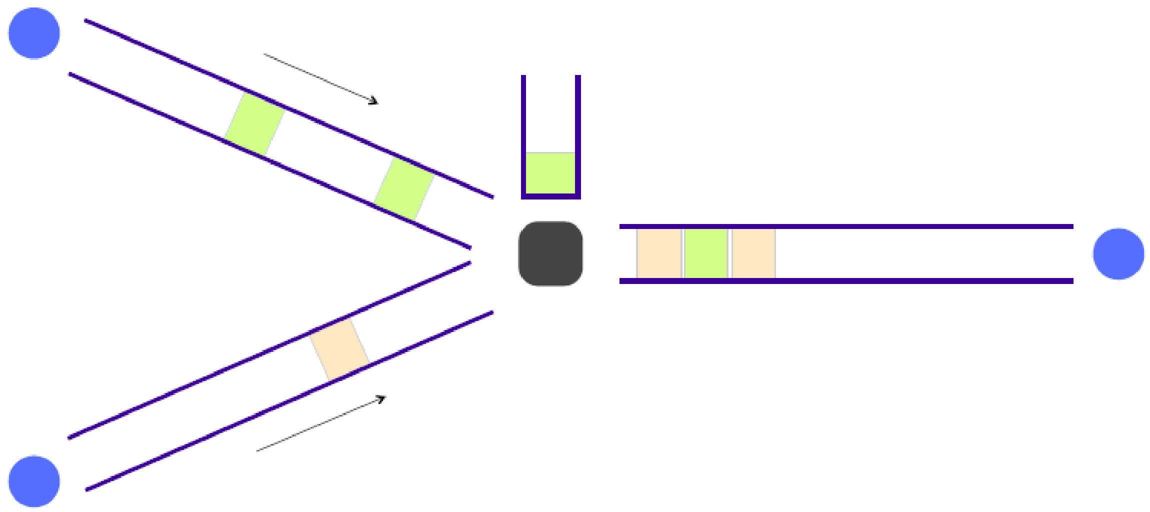


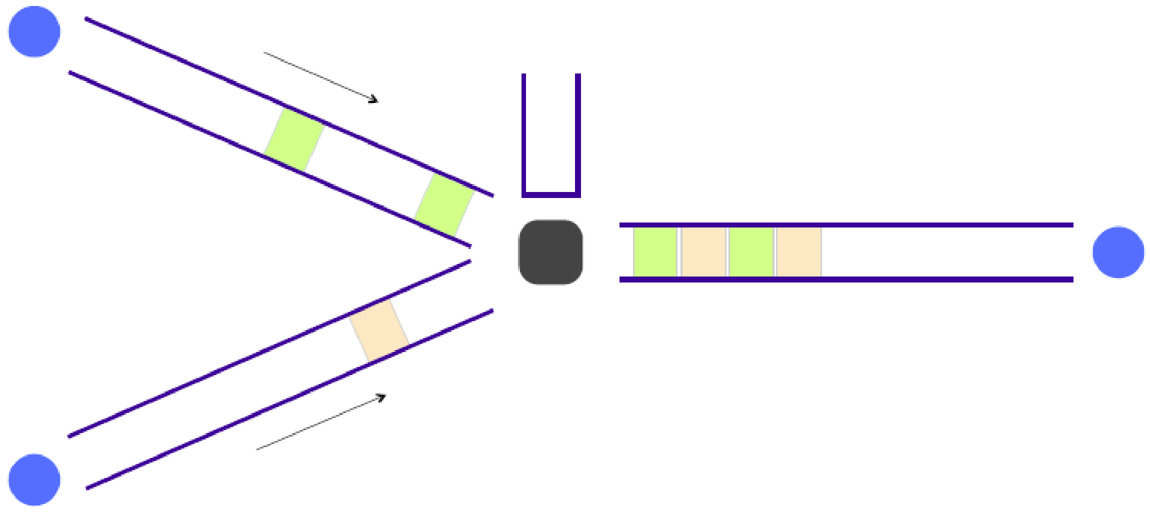

No overload


Queue
Transient overload!

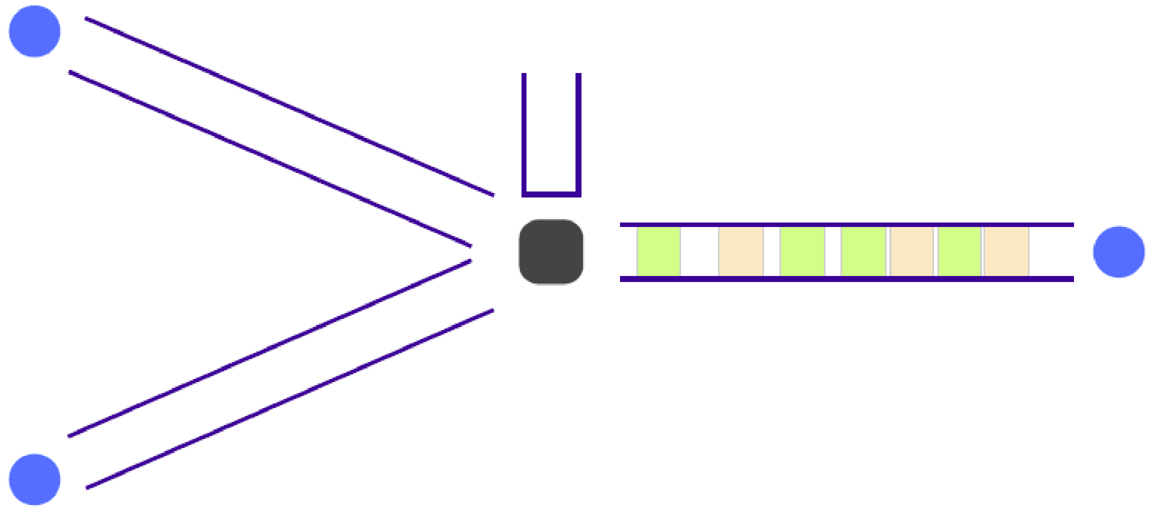

Transient overload!







Queues absorb transient bursts, but introduce queueing delays



The time a packet has to sit in a buffer before being processed depends on the traffic pattern

Queueing delay depends on:

arrival rate at the queue

transmission rate of the outgoing link

traffic **burstiness**

average packet arrival rate	a	[packet/sec]
transmission rate of outgoing link	R	[bit/sec]
fixed packets length	L	[bit]
average bits arrival rate	La	[bit/sec]
traffic intensity	La/R	

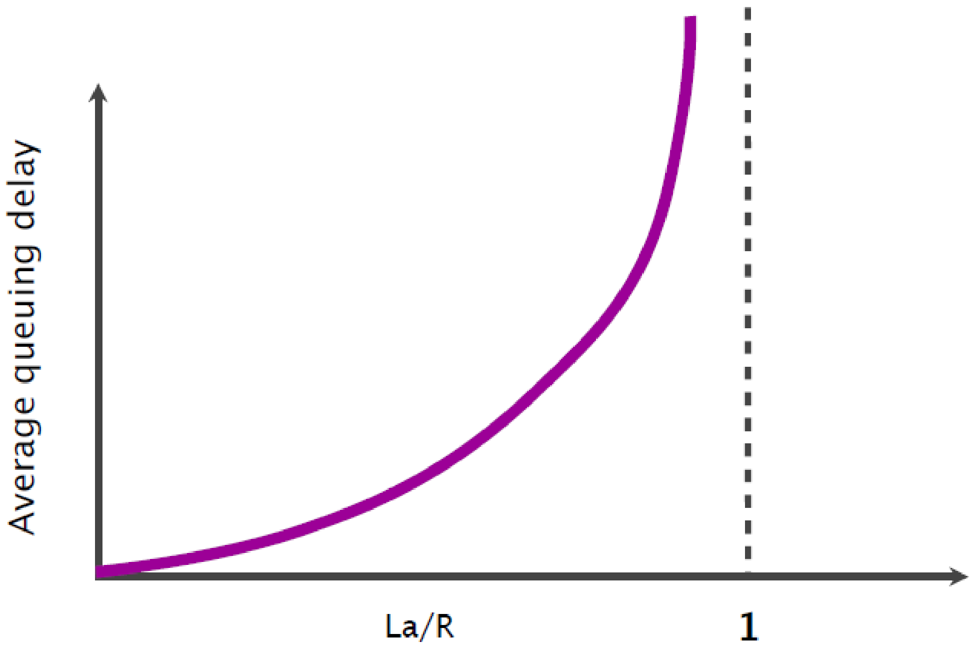
When the traffic intensity is >1, the queue will increase without bound, and so does the queuing delay

Golden rule

Design your queuing system,

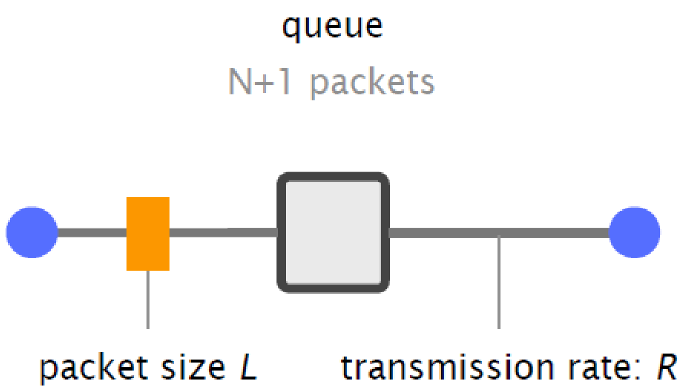
so that it operates far from that point

When the traffic intensity is <=1, queueing delay depends on the burst size



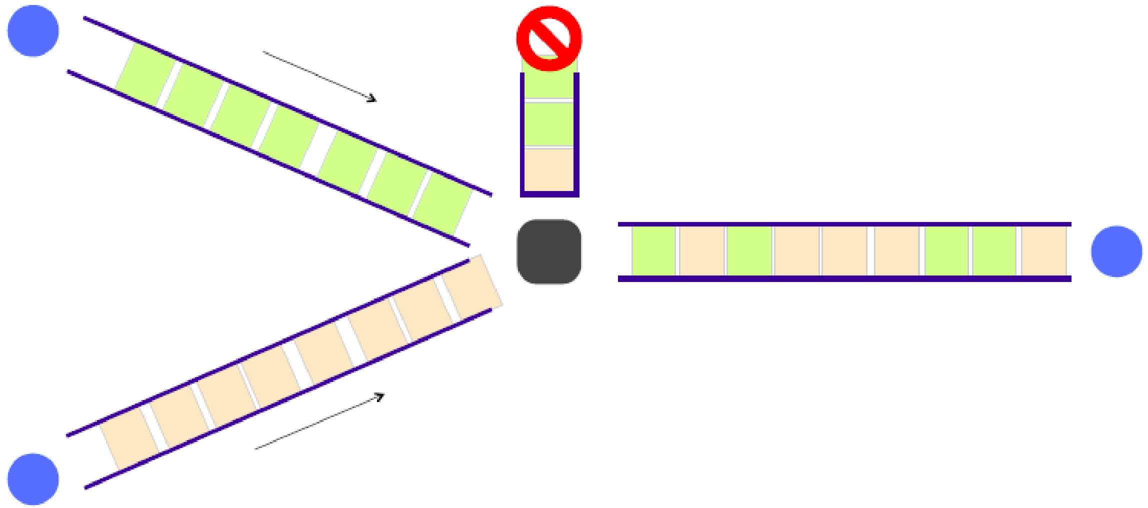
Loss

In practice, queues are not infinite.
There is an upper bound on queuing delay.



queuing delay upper bound: N*L/R

If the queue is persistently overloaded, it will eventually drop packets (loss)



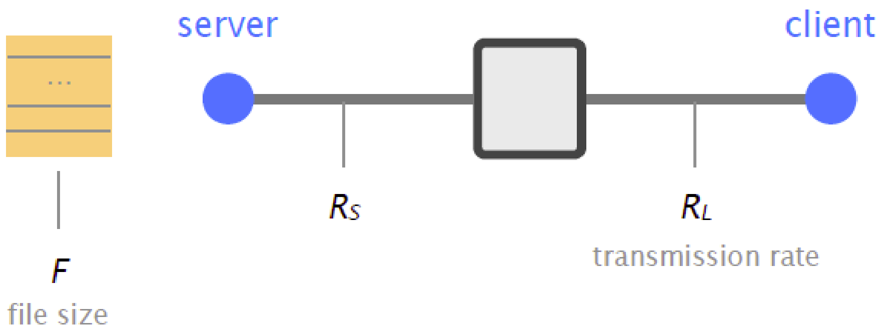
Throughput

throughput

The throughput is the instantaneous rate at which a host receives data

$$\text{Average throughput [\#bits/sec]} = \frac{\text{data size} \quad \text{[\#bits]}}{\text{transfer time} \quad \text{[sec]}}$$

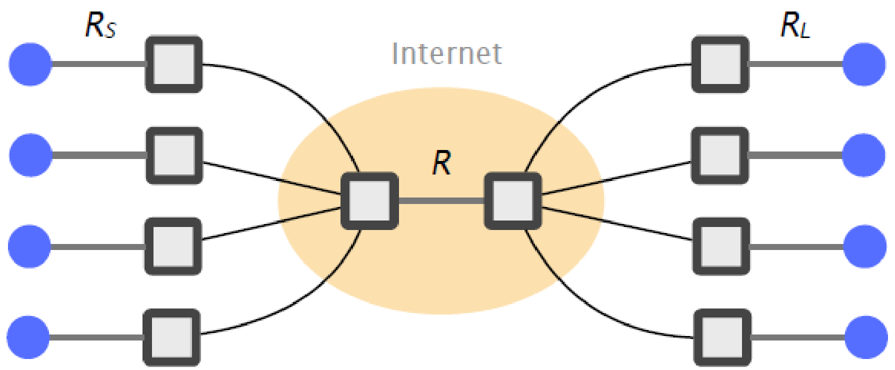
To compute throughput, one has to consider the bottleneck link



Average throughput $\min(R_S, R_L)$

= transmission rate of the bottleneck link

To compute throughput, one has to consider the bottleneck link… and the intervening traffic

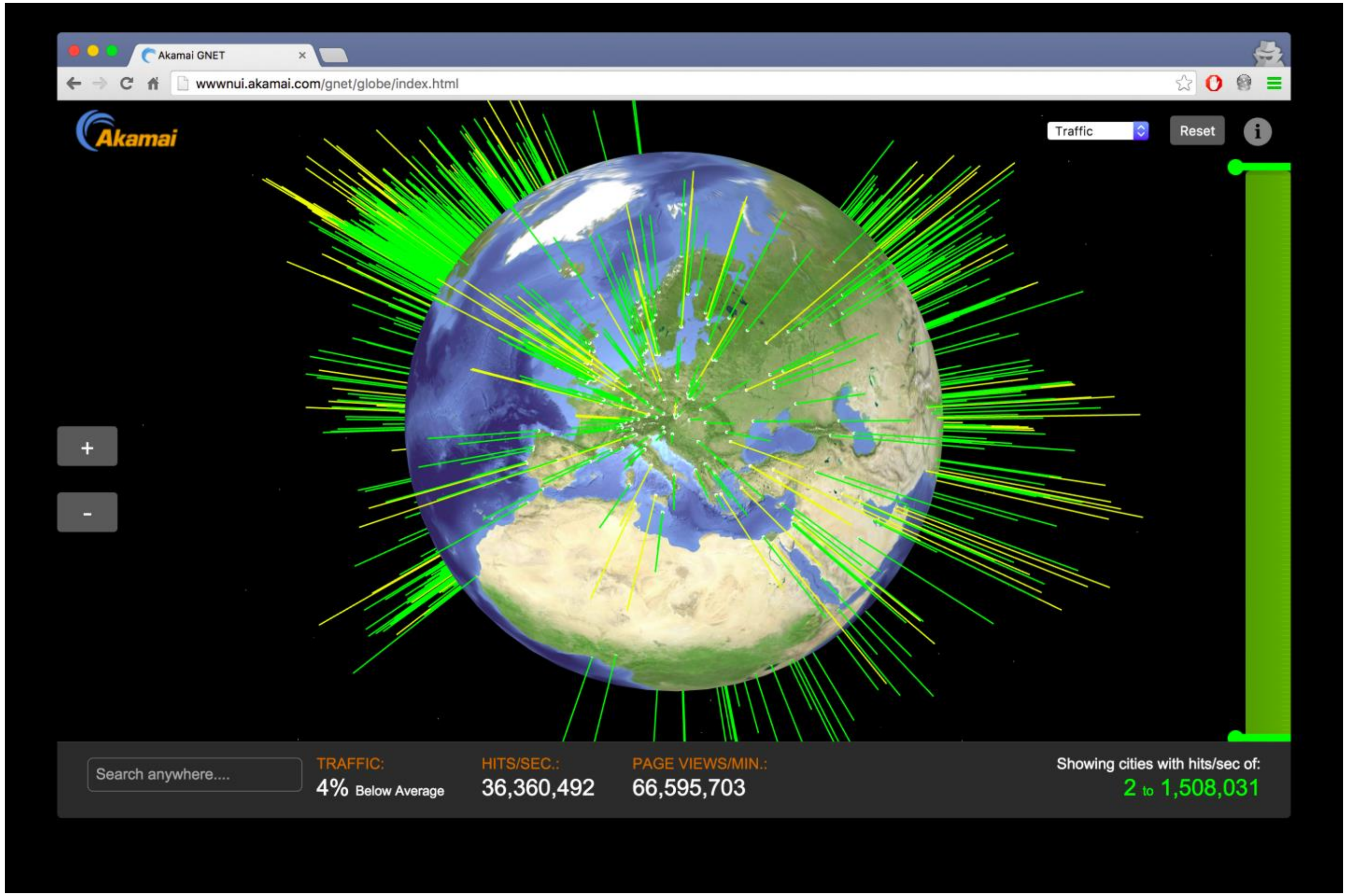


if $4 * \min(R_S, R_L) > R$ — the bottleneck is now in the core, providing each download R/4 of throughput

As technology improves, throughput increase & delays are getting lower except for propagation

(speed of light)

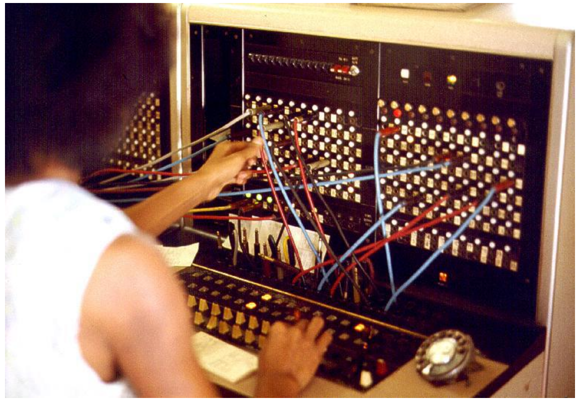
Because of propagation delays,
Content Delivery Networks move content closer to you



* http://wwwnui.akamai.com/gnet/globe/index.html

Brief history of Internet

The Internet history starts in the late 50's…



Phone networks = the communication network

fully circuit-switched

People wanted to use networks for other things

defense, computers, etc.

circuit-switching does not fit to these new requirements…

inefficient for bursty loads and not resilient

Three main questions

Paul Baran RAND

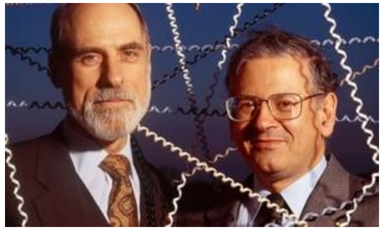
How can we design a more resilient network?

… led to the invention of packet switching

Leonard Kleinrock
UCLA

How can we design a more efficient network?

… also led to the invention of packet switching



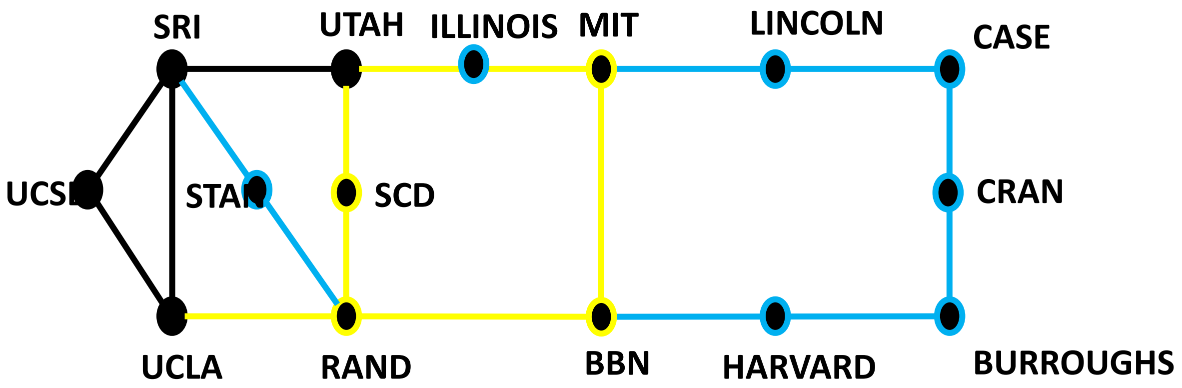
Vint Cerf & Bob Kahn
DARPA

How can we connect all these networks together?

… the invention of Internet as we know it

The 60's was all about packet switching…

Advanced Research Projects Agency NETwork (ARPANET)



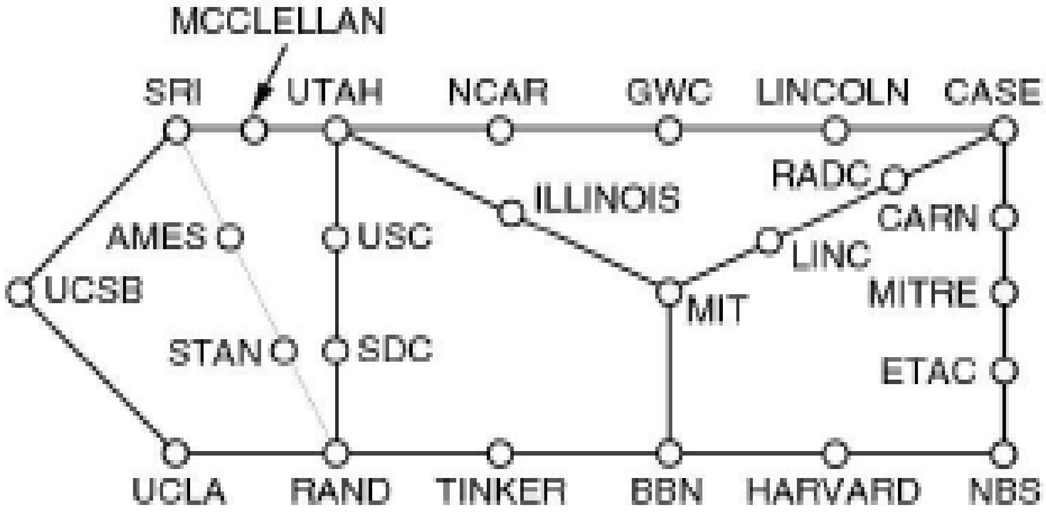
1969 december

1970 july

1971 march

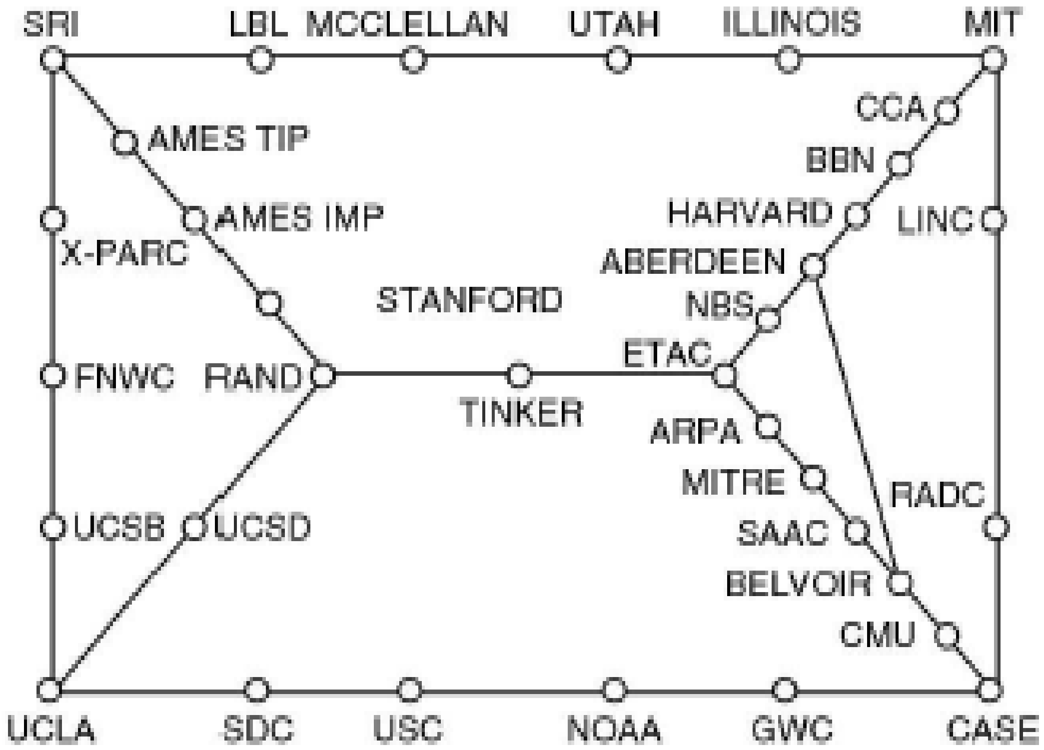
ARPANET

April 1972

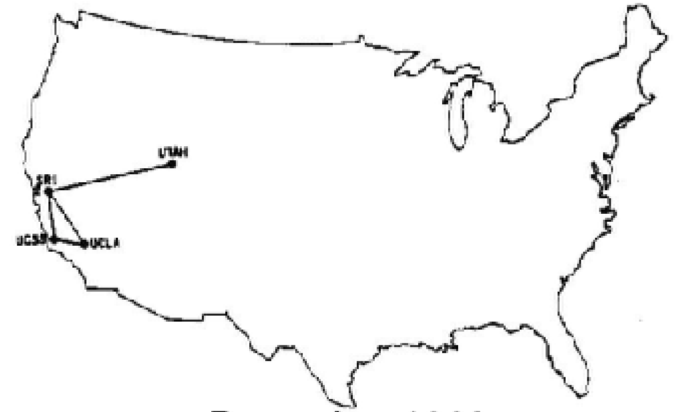


ARPANET

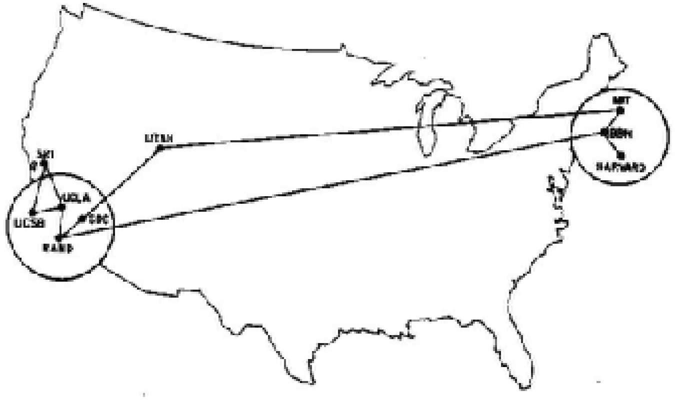
September 1972



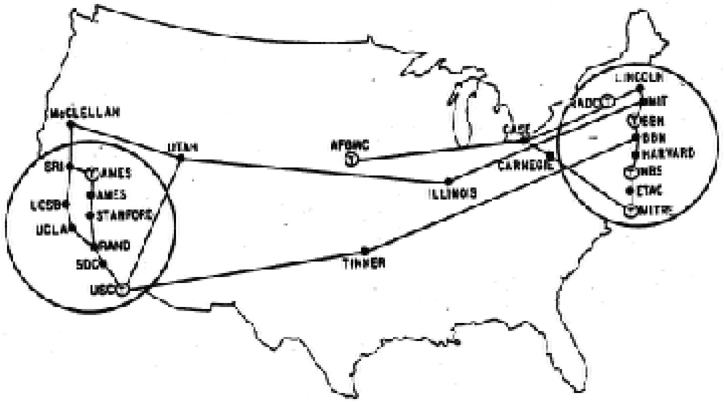
ARPANET



Dezember 1969



Juni 1970



März 1972

Juli 1977

The first message over the Internet : **„LO”**

29. Oct. 1969

Leonard Kleinrock from UCLA tries to log in a Stanford computer

UCLA We typed the L… Do you see it?

Yes! We see the L Stanford

We typed the O… Do you see it?

Yes! We see the O

We typed the G.

… and the system crashed …

* http://ftp.cs.ucla.edu/csd/first_words.html

The 70’s about Ethernet, TCP/IP and email...

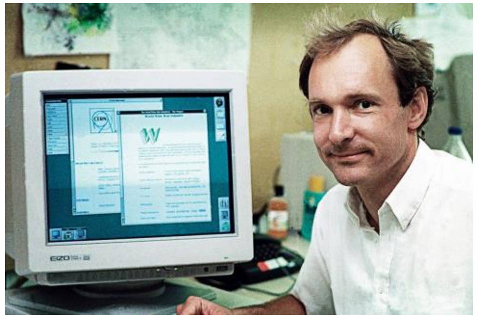
1971	**Network Control Program (NCP)** Predecessor of TCP/IP
1972	**Email and Telnet**
1973	**Ethernet**
1974	**TCP/IP** Paper of Vint Cerf and Bob Kahn

80’s when TCP/IP went mainstream

1983	**NCP to TCP/IP**
	Domain Name Service (DNS)
1985	**NSFNet (TCP/IP)**
198x	**First Internet crashes caused by congestion**
1986	**Van Jacobson saves the Internet**
	congestion control

Van Jacobson

90’s – the Internet going commercial…



1989	**ARPANET closed** **Birth of the WEB** Tim Berners Lee (CERN)
1993	**First search engine (Excite)**
1995	**NSFNet closed**
1998	**Google reinvents searching**

The new millenium bringing Web 2.0

1998	**IPv6 standardization**
2004	**Facebook goes online**
2006	**Google buys YouTube**
2007	**Netflix strats streaming videos**
2007	**First iPhone with mobile Internet access**

Fast Internet access everywhere, every device needs an Internet connection

2009	**Mining of the Bitcoin genesis block**
	Fast mobile Internet access: 4G/LTE
	IoT boom Internet of Everything
2018	**Only 26% of Alexa Top 1000 sites reachable over IPv6**
Soon?	**Encrypted transport protocols** For example QUIC **Ultra-fast mobile access – 5G**

Standardization of Internet protcols

I E T F

IRTF

They uses collective humming to get consesus for technical discussions…

https://twitter.com/biellacoleman/status/1021138393252712449?lang=en

To be continued…