The Internet

The network of networks

Outline

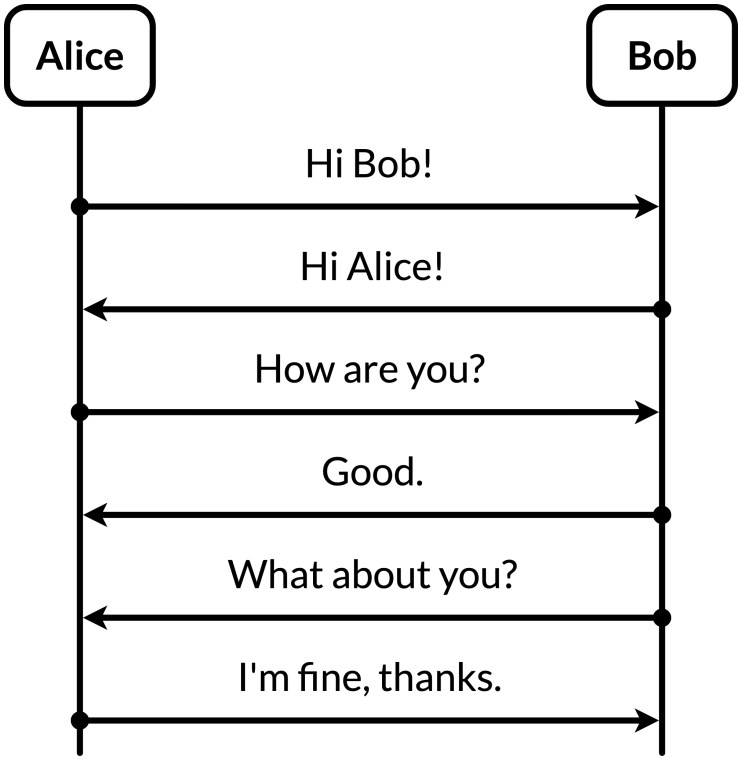
- Communication protocol with error recovery
- Signal routing and packet switching
- Internet protocol and Internet layers
- IP addresses, latency, and port numbers
- TCP, TLS, DNS, and HTTP

Slides available at it-course.ch/Internet.pdf

I wrote an introduction at ef1p.com/internet

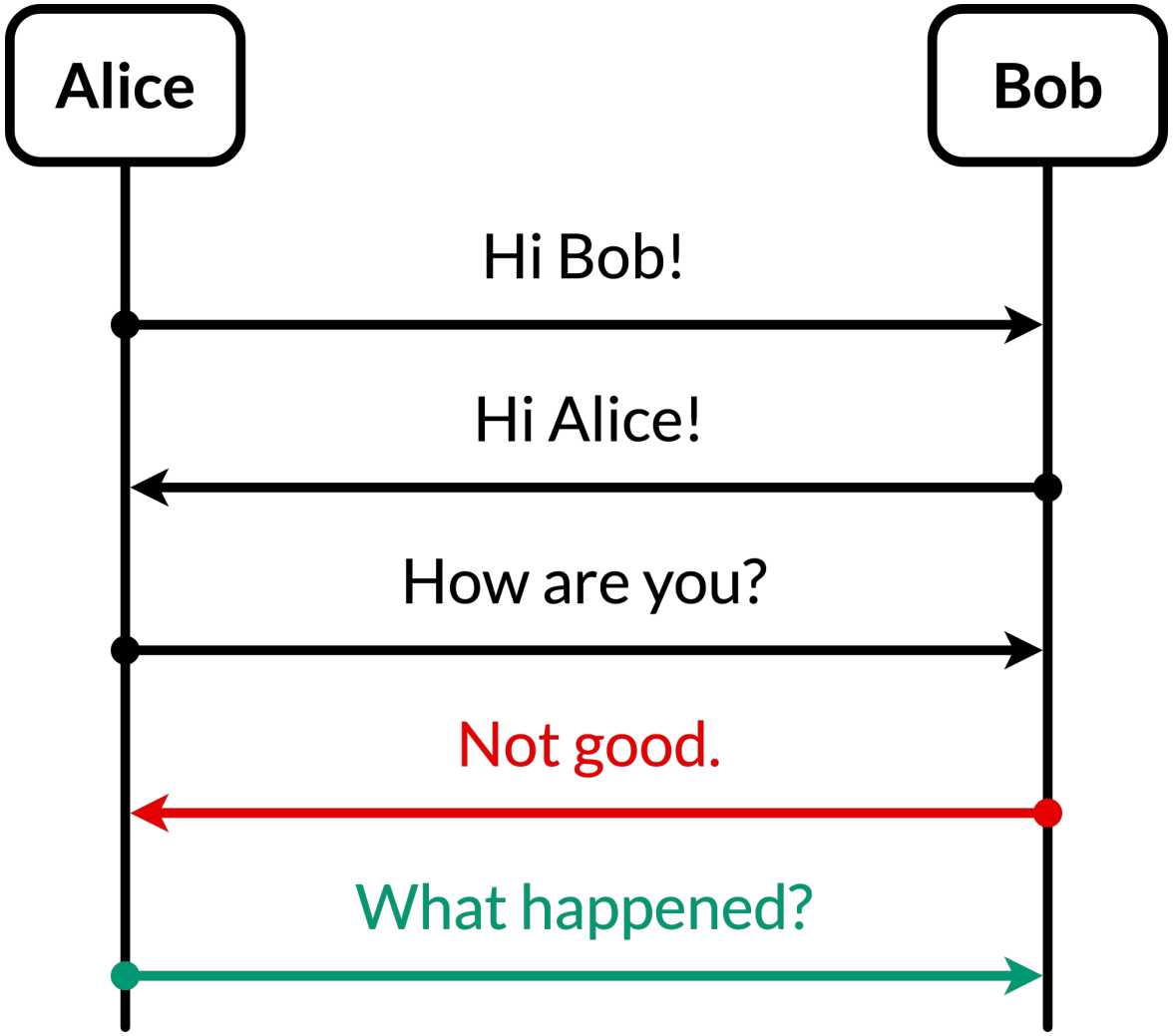
Communication protocol

Can often be displayed with a sequence diagram.

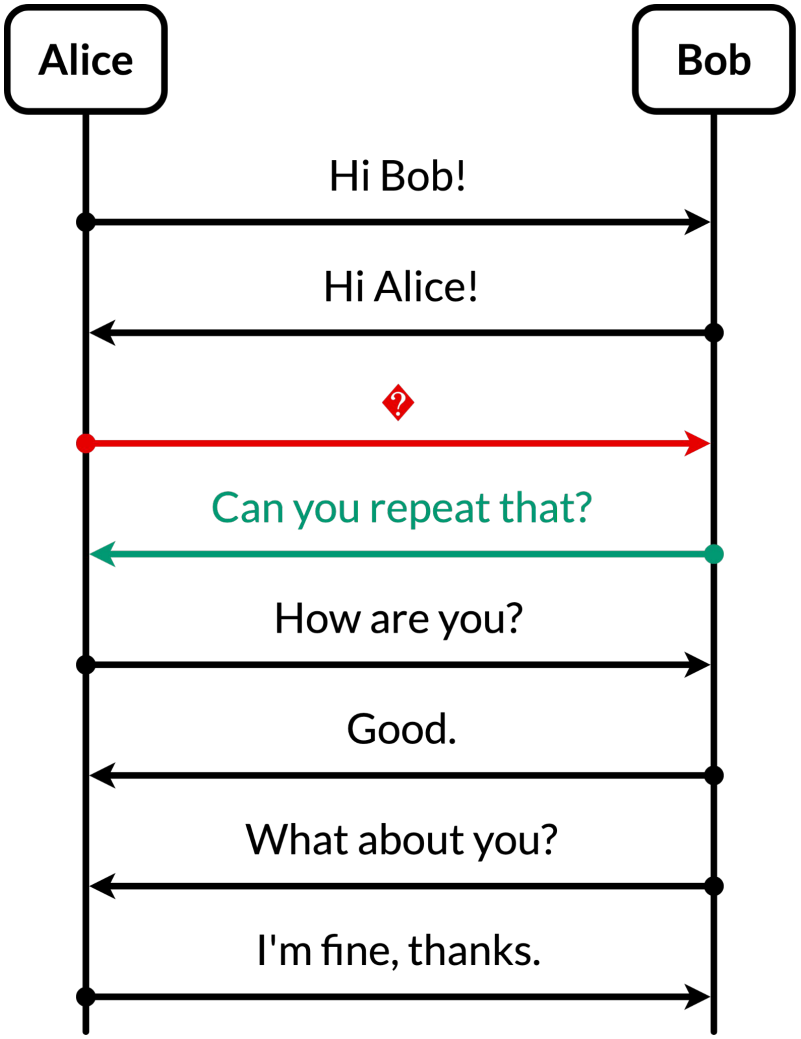


Protocols should be able to recover from errors.

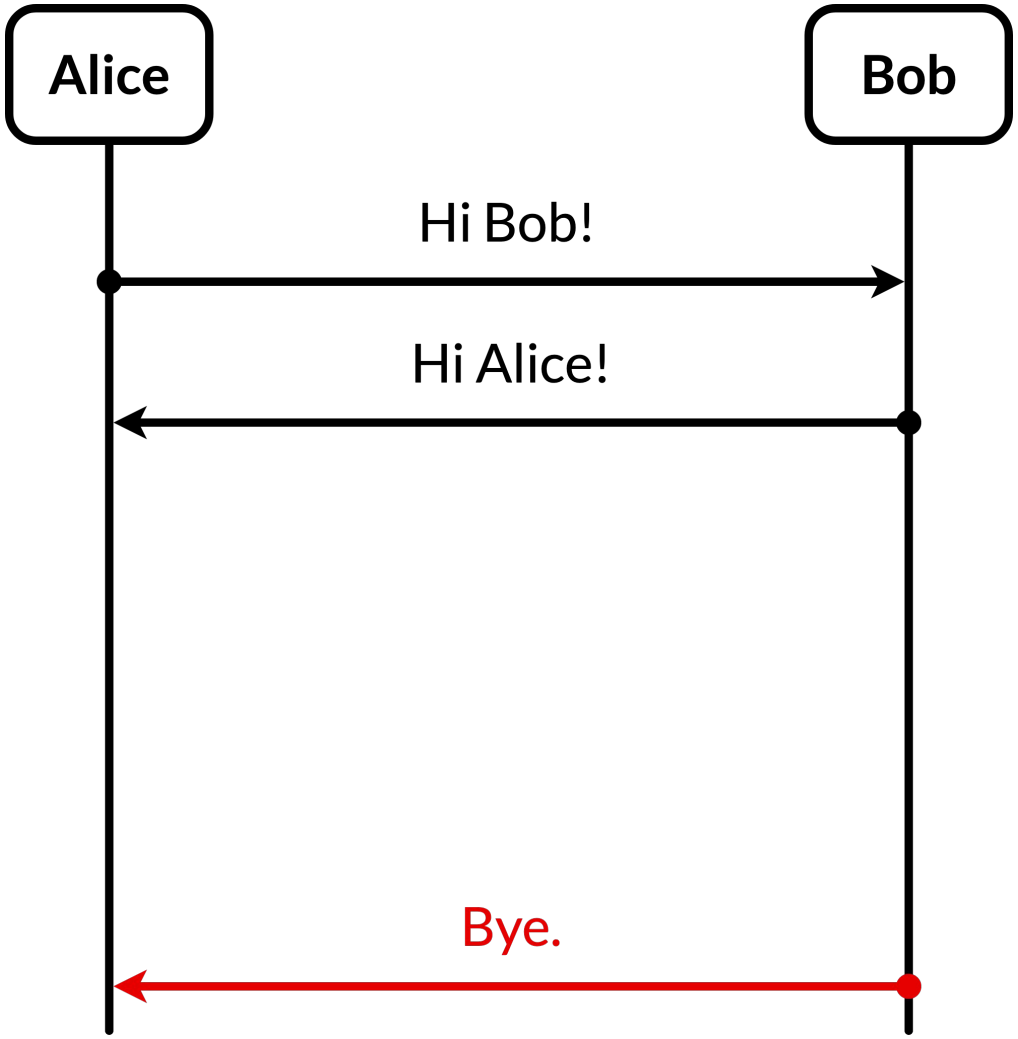
Example: protocol deviation



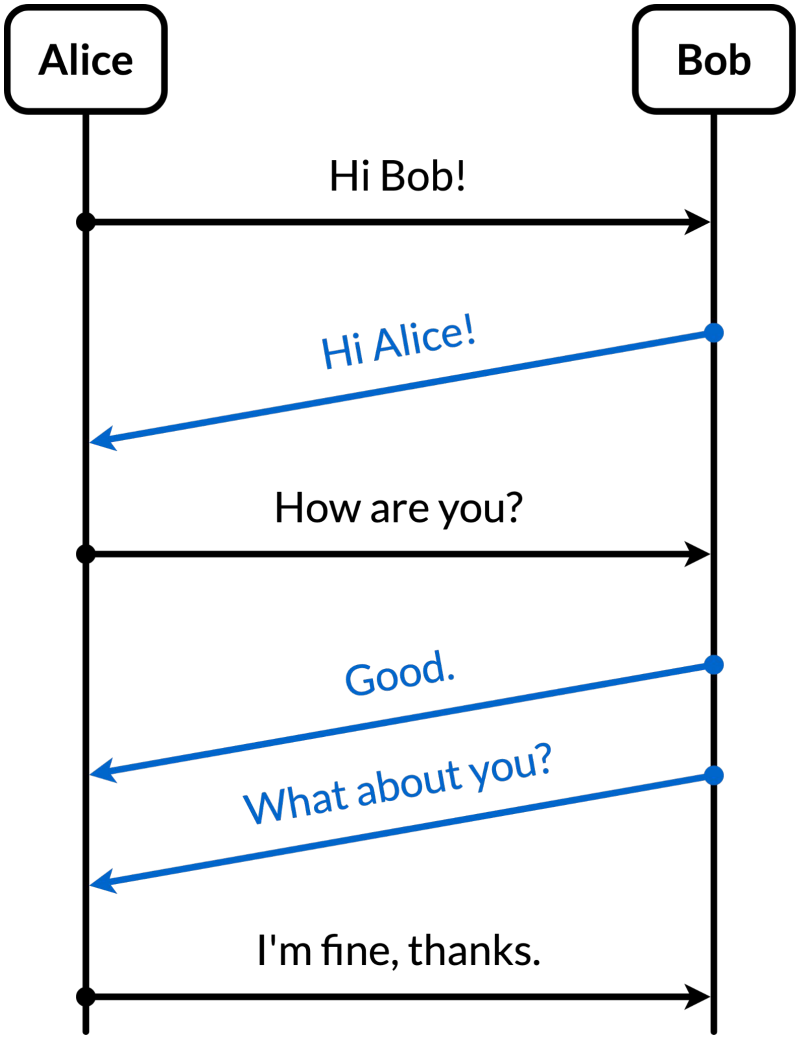
Example: data corruption



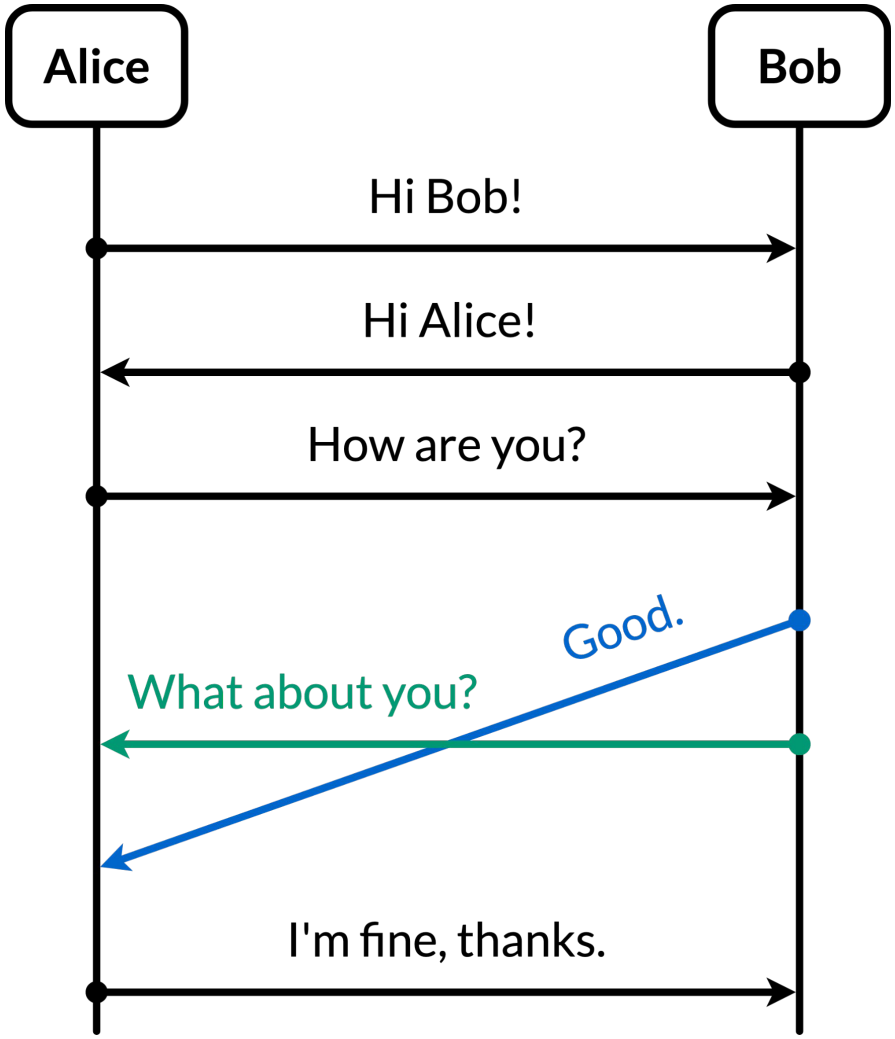
Example: connection loss



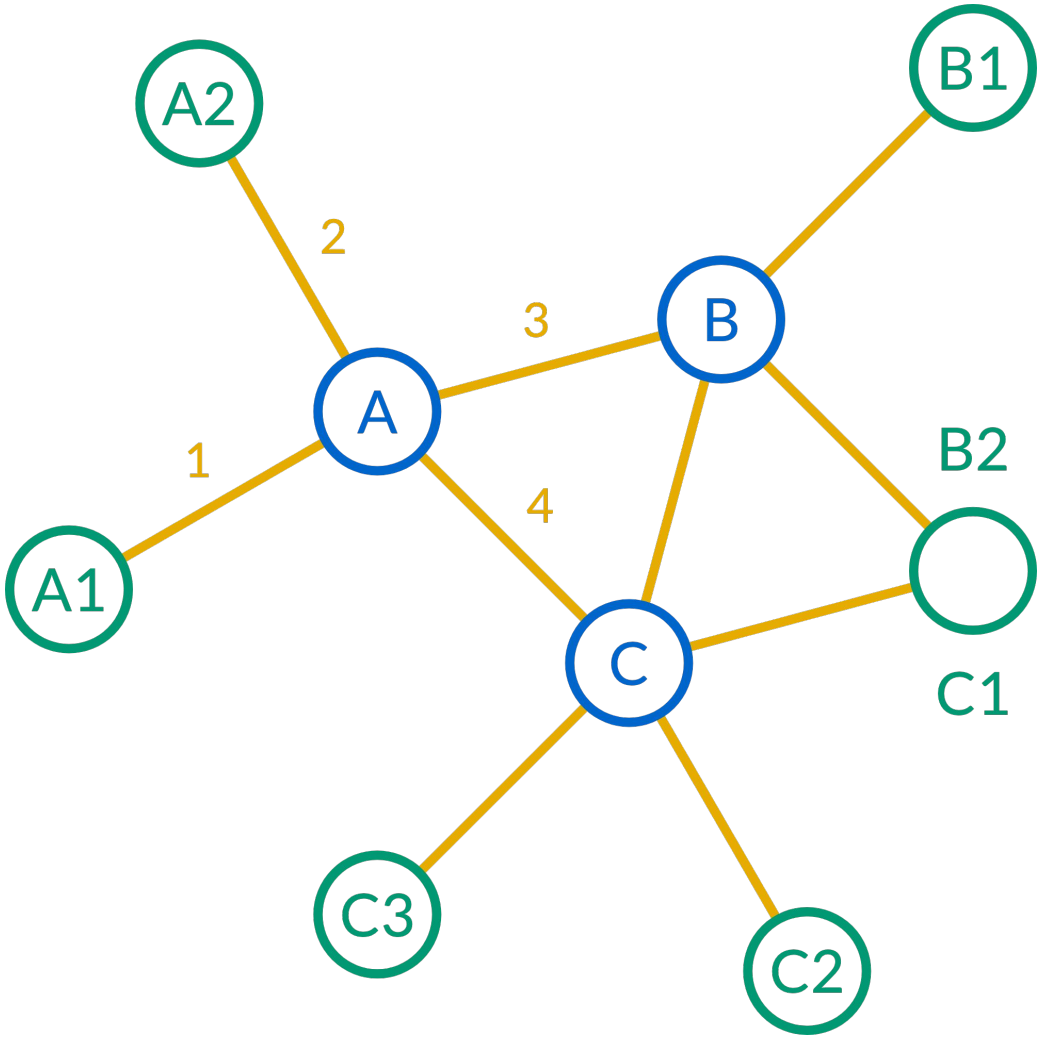
Example: network latency



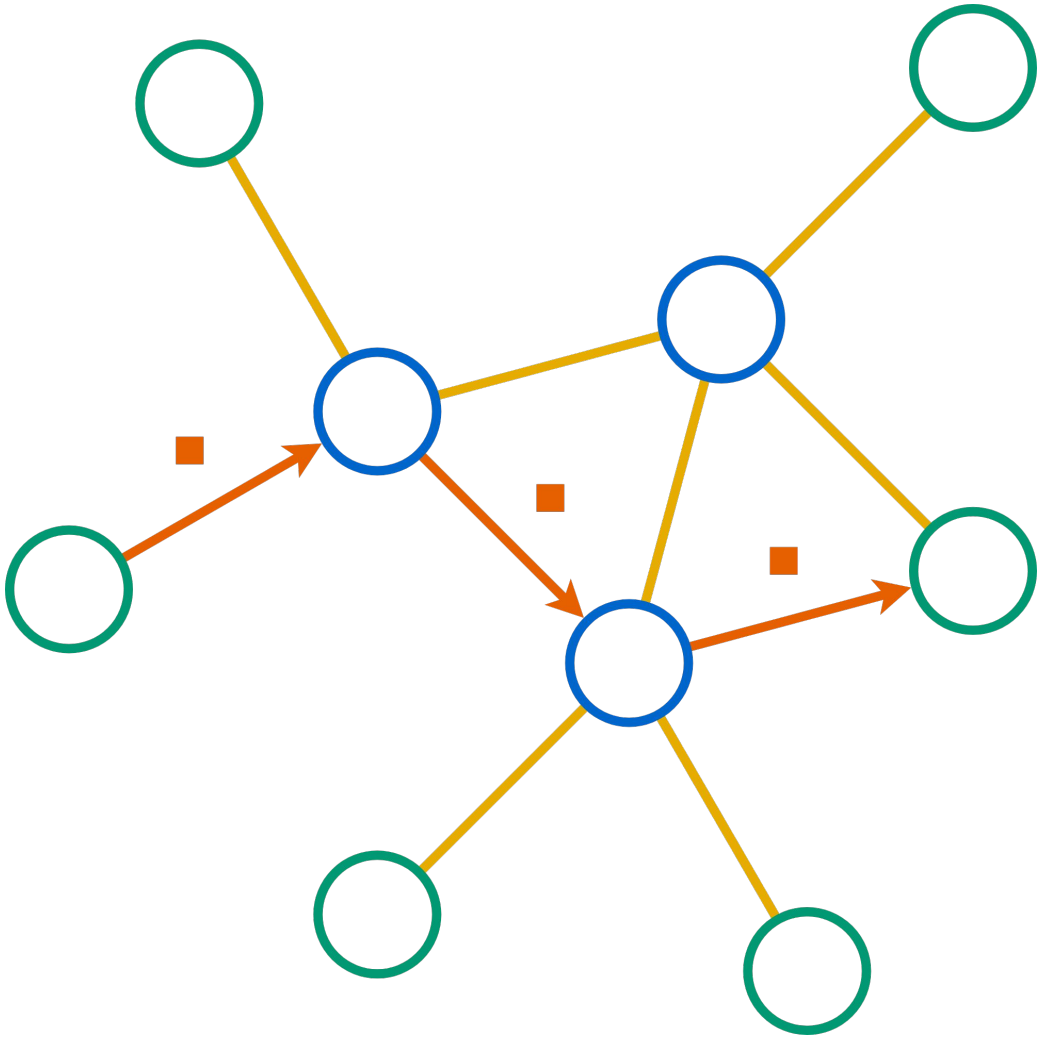
Example: out-of-order delivery



Signal routing



Packet switching



The Internet Protocol (IP)

The Internet Protocol made independent networks compatible by introducing a common packet and address format, allowing routing across networks.

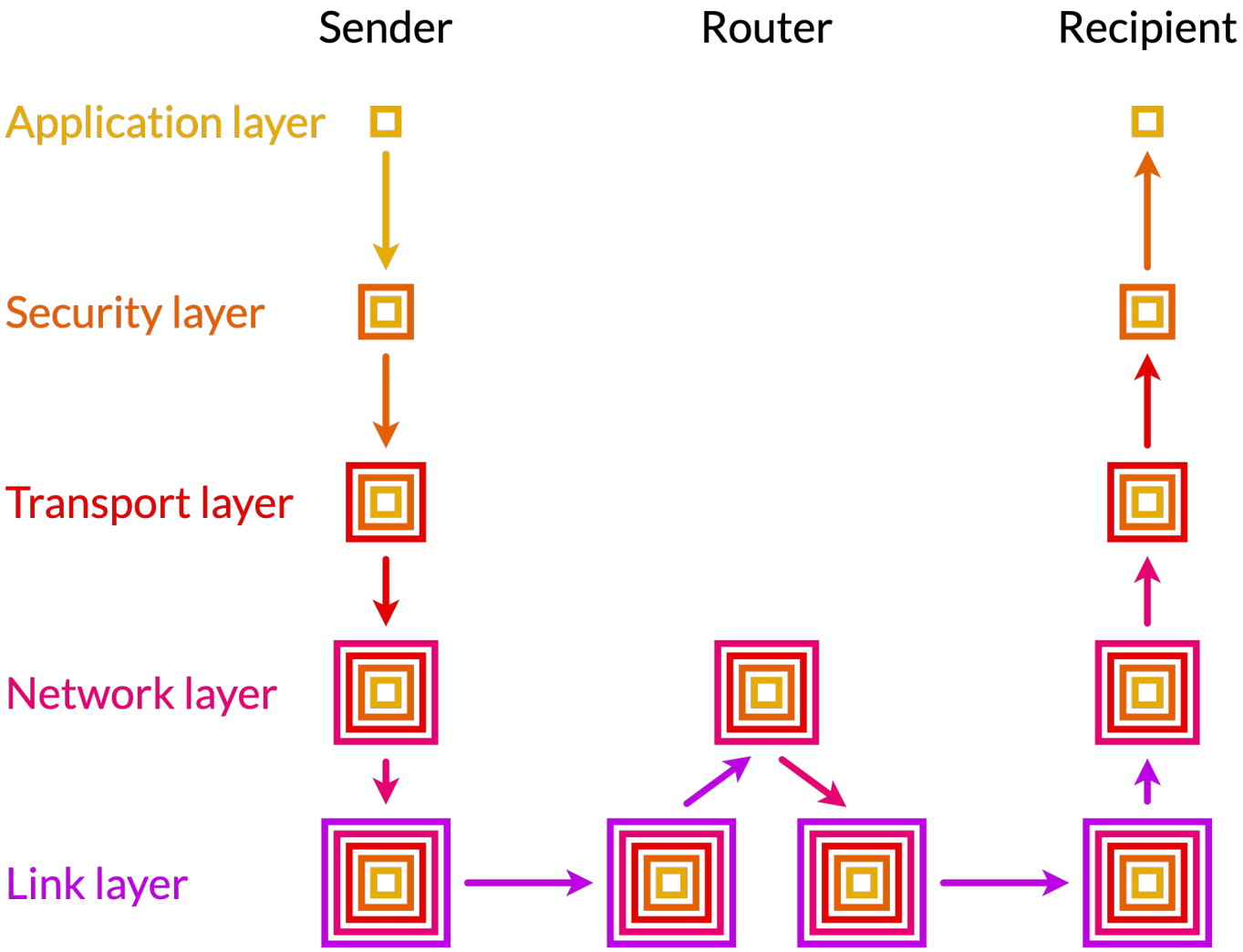
Communication over the Internet is unreliable: Packets can get lost or arrive out of order.

Internet layers

The Internet operates in layers to be more flexible.

Name	Purpose	Example
Application layer	Application logic	HTTP
Security layer	Encryption and authentication	TLS
Transport layer	Typically reliable data transfer	TCP
Network layer	Packet routing across the Internet	IP
Link layer	Handling of the physical medium	Wi-Fi

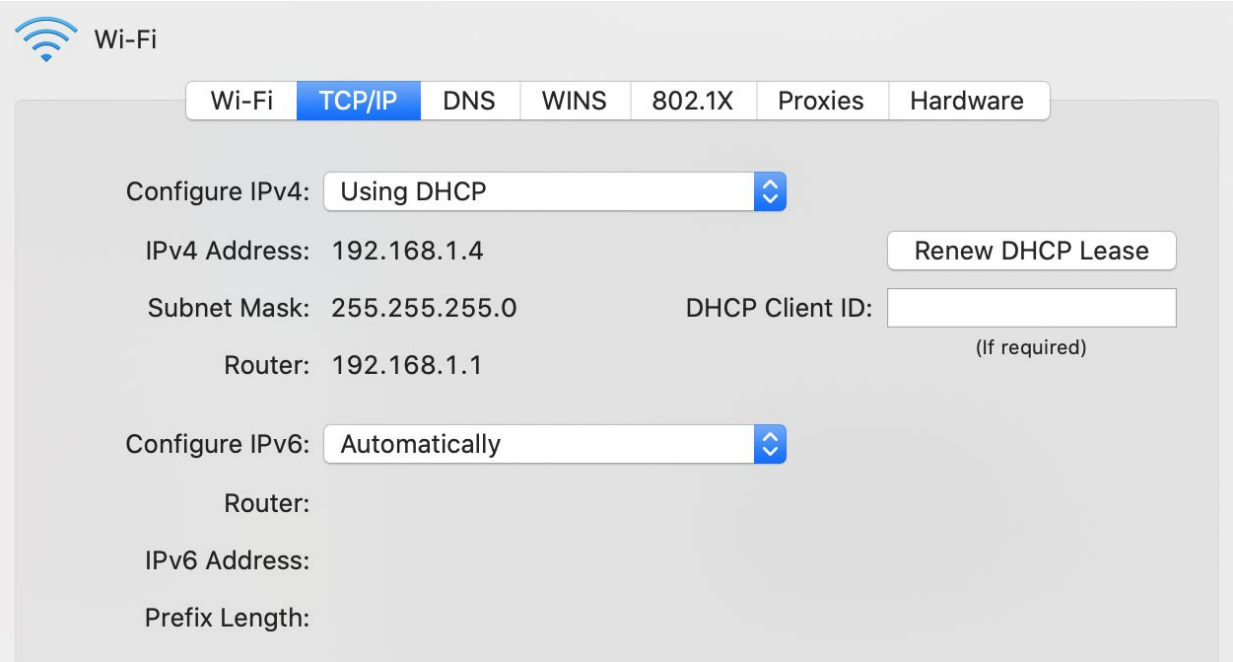
Internet layers visualized



IP addresses

IP(v4) address: 32-bit identifier for netw. interface

Often local address ≠ global address due to network address translation (NAT) done by router



IP geolocation

The Internet is also a physical network.

IP addresses assigned by subnetworks.

Can be used to locate users approximately.

Demonstrate this with this tool.

Network latency

Determine the round-trip time with ping.

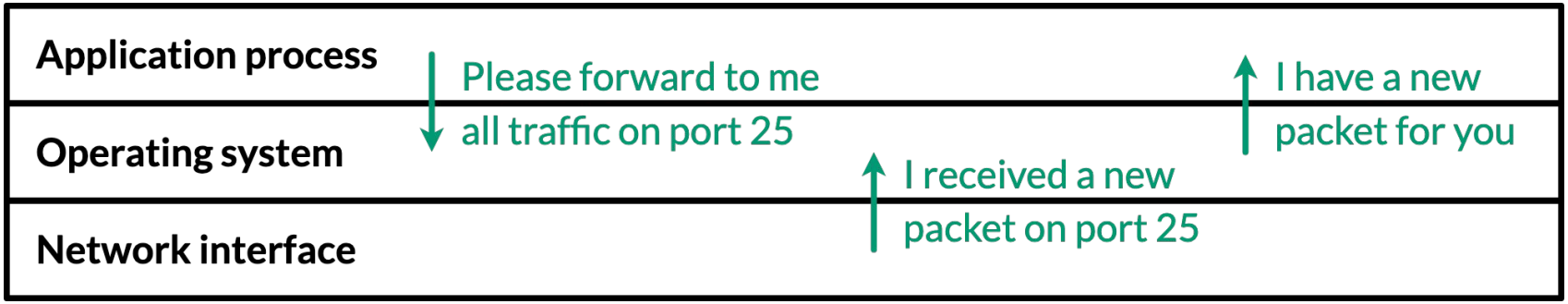
```
ping -c 5 ethz.ch
ping -c 5 google.com
ping -c 5 example.com
ping -c 5 ietf.org
```

(Ashburn near Washington DC)

(San Jose near San Francisco)

Port numbers

The IP address identifies a computer, whereas the port number identifies a process on the computer.



Client-server model

A client requests a service from a server:

The server's port number depends on the service; the client's port number can be chosen randomly.

Wikipedia has a list of established port numbers.

Transmission Control Protocol (TCP)

The Transmission Control Protocol (TCP) provides in-order data transfer between two computers.

The sender and the receiver buffer all packets.

The receiver reorders the incoming packets based on a sequence number and asks the sender to resend any missing packets.

TCP also provides flow and congestion control.

Transport Layer Security (TLS)

Transport Layer Security (TLS) is the main protocol to provide confidential and authenticated communication over the Internet.

It uses TCP on the transport layer and provides:

- Party authentication with public-key certificates,
- Confidentiality w. symmetric-key cryptography,
- Message authentication with a hash function.

Domain Name System (DNS)

The Domain Name System (DNS) is a hierarchical namespace of easily memorizable domain names and an application-layer protocol to access public information associated with such names.

It is most commonly used to look up the IP address of a server in order to connect to it.

You can register a domain name at a registrar.

Example: `dig +short ef1p.com`

HyperText Transfer Protocol (HTTP)

HTTP is the application-layer protocol of the World Wide Web (www) to transfer files over the Internet.

It's a text-based protocol with two versions:

- HTTP over TCP on port 80 and
- HTTPS over TLS on port 443.

```
$ openssl s_client -quiet -crlf -connect
explained-from-first-principles.com:443
GET /internet/ HTTP/1.0
Host: explained-from-first-principles.com
```

Uniform Resource Locator (URL)

A Uniform Resource Locator (URL) identifies a resource with the following syntax:

scheme://domain:port/path?querystring#fragment

Example: https://ef1p.com/internet/#preface

Web languages

A web browser retrieves a web page from a web server and renders its content for you. The server can send static files or generate them dynamically.

There are three languages for web pages:

- HyperText Markup Language (HTML) for content,
- Cascading Style Sheets (CSS) for design,
- JavaScript (JS) for interactivity.

Only JavaScript is a real programming language.

HyperText Markup Language (HTML)

HTML structures the content of a web page with tags. The content is put between an opening tag, e.g. `<p>`, and a corresponding closing tag, `</p>`.

Tags can have attributes: `<a href="d.html">`.

Elements can be nested: `<p>A <i>B</i></p>`.

… unless they're void: `<img src="me.png"/>`.

There are many HTML tags, but you don't have to know them all. Just google what you're looking for.

HTML example

```
<!doctype html>
<html>
 <head>
  <title>Title of web page</title>
 </head>
 <body>
  <h1>Heading</h1>
  <p>Paragraph with a
    <a href="d.html">link</a>.</p>
  <img src="image.jpg"/>
 </body>
</html>
```