



Why Should I Care About Networking?

Ivan Pepelnjak (ip@ipSpace.net)
Network Architect

ipSpace.net AG



Who is Ivan Pepelnjak (@ioshints)

Past

- Kernel programmer, network OS and web developer
- Sysadmin, database admin, network engineer, CCIE
- Trainer, course developer, curriculum architect
- Team lead, CTO, business owner



Present

- Network architect, consultant, blogger, webinar and book author

Focus

- Large-scale data centers, clouds and network virtualization
- Scalable application design
- Core IP routing/MPLS, IPv6, VPN



IT ALL WORKED GREAT

ON MY LAPTOP

Fallacies of Distributed Computing

- The network is reliable.
- Latency is zero.
- Bandwidth is infinite.
- The network is secure.
- Topology doesn't change.
- There is one administrator.
- Transport cost is zero.
- The network is homogeneous.

Peter Deutsch (1994)

Also known as: I Believe in Bandwidth Fairy



1 SECOND CLICK-TO-SCREEN

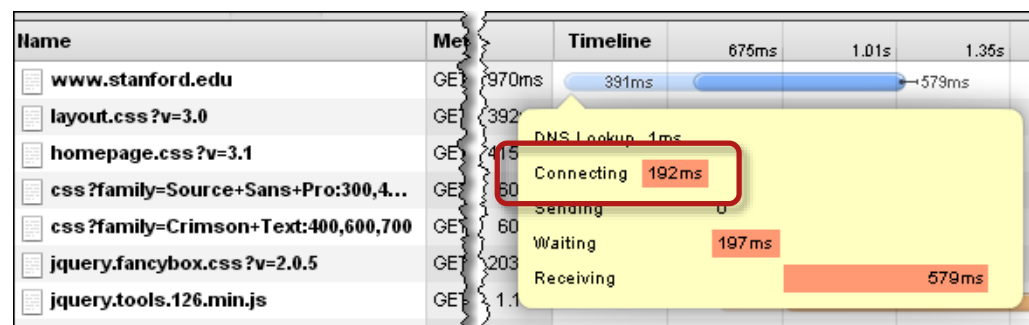
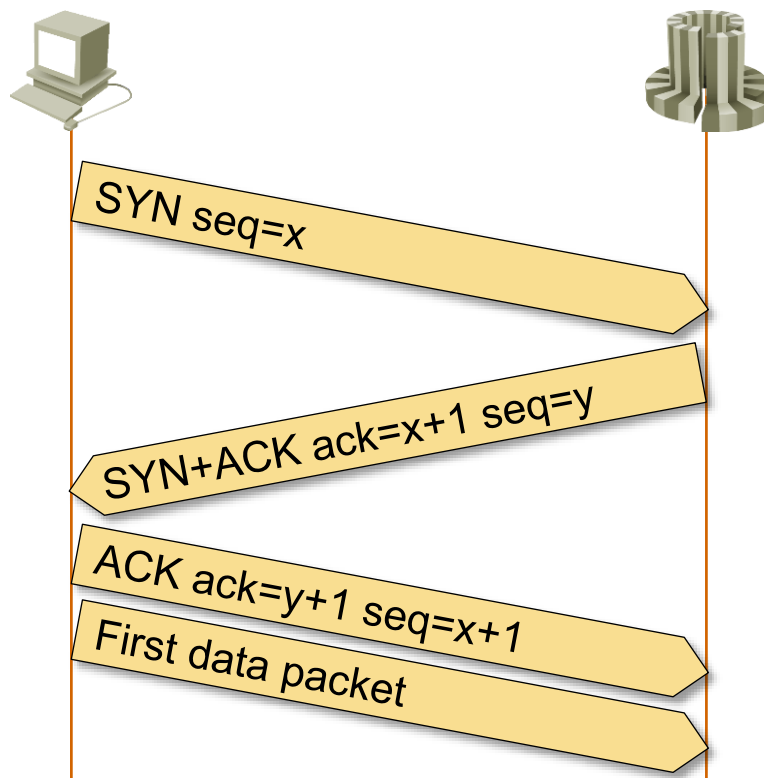
YEAH, NO BIG DEAL



1 SECOND CLICK-TO-SCREEN

YEAH, NO BIG DEAL

This Is Why You Need the Details



- TCP session established with a 3-way handshake
- RTT delay before first user data is sent

Does It Really Matter? We're Not in Antarctica

```
$ ping www.nil.com
```

```
Pinging www.nil.com [192.168.253.10] with 32 bytes of data:  
Reply from 192.168.253.10: bytes=32 time=8ms TTL=253  
Reply from 192.168.253.10: bytes=32 time=8ms TTL=253  
Reply from 192.168.253.10: bytes=32 time=8ms TTL=253  
Reply from 192.168.253.10: bytes=32 time=9ms TTL=253
```

Fiber Internet access

```
$ ping www.nil.com
```

```
Pinging www.nil.com [193.110.145.49] with 32 bytes of data:  
Reply from 193.110.145.49: bytes=32 time=369ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=282ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=409ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=267ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=242ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=223ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=178ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=167ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=193ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=136ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=249ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=228ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=193ms TTL=244  
Reply from 193.110.145.49: bytes=32 time=167ms TTL=244
```

**3G mobile access
over Bluetooth**

Remember: Latency is never zero. It could be higher than expected

1 Second Click-to-Screen Is Extremely Hard

*What Every Web Developer Should Know About
Networking and Browser Performance*



High Performance

Browser Networking

O'REILLY®

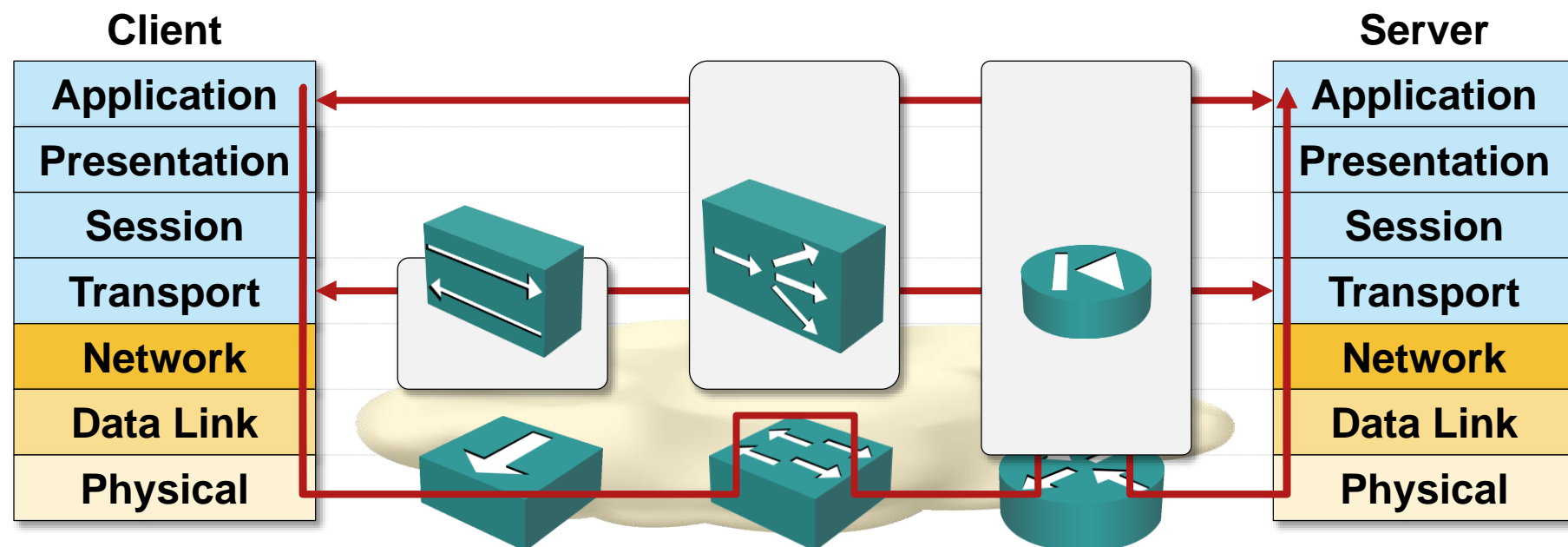
Ilya Grigorik



7-LAYER OSI MODEL?

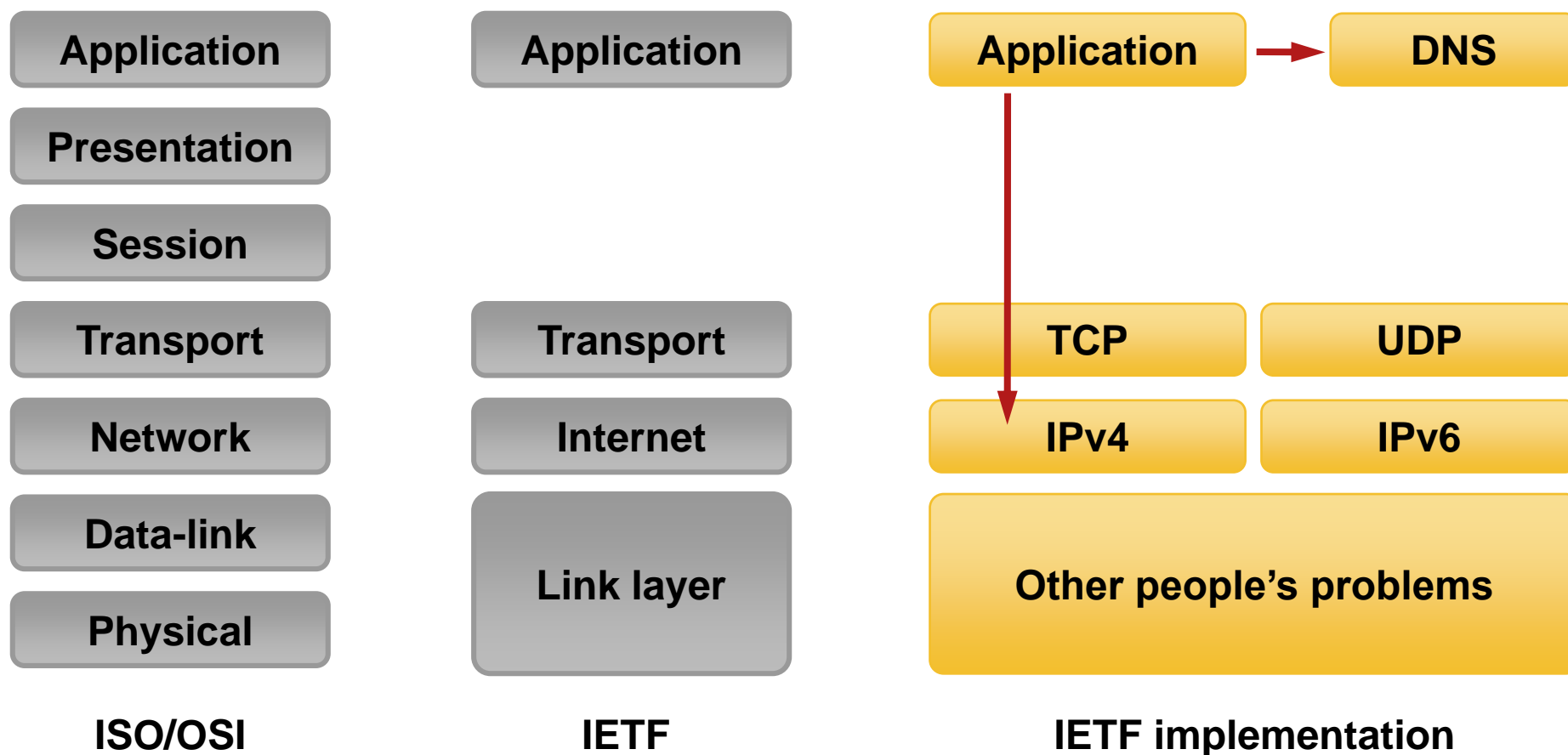
**I DON'T HAVE
TIME FOR THAT**

Data Communications and Onions



Things Like NAT and Firewalls Will Trip You Up

TCP/IP: Broken By Design



- Session layer is missing (session endpoints tied to IP addresses)
- IP addresses are visible to applications (no L3-L7 abstraction)
- DNS is an optional add-on application

Socket API: Broken By Design

Ideal

```
conn = Network.Connect("example.com", "http")
```

TBD

OK

```
conn = new Socket("example.com", 80)
```

Java

Broken

```
memset(&hints, 0, sizeof(hints));
hints.ai_family = PF_UNSPEC;
hints.ai_socktype = SOCK_STREAM;
error = getaddrinfo("example.com", "http", &hints, &res0);
if (error) { errx(1, "%s", gai_strerror(error)); }

s = -1;
for (res = res0; res; res = res->ai_next) {
    s = socket(res->ai_family, res->ai_socktype, res->ai_protocol);
    if (s < 0) { cause = "socket"; continue; }

    if (connect(s, res->ai_addr, res->ai_addrlen) < 0) {
        cause = "connect";
        close(s);
        s = -1;
        continue;
    }

    break; /* okay we got one */
}
if (s < 0) { err(1, "%s", cause); }
```

Socket API

Consequences of Broken TCP/IP Stack & API

- Every application reinvents the wheel (sometimes badly)
- New network/transport protocols are a royal pain (see: IPv6) and require application changes
- Endpoint mobility is hard to implement (network-layer addresses have to move)
- No automatic transport/network protocol selection
- No load balancing across multiple endpoint addresses
- No automatic session reconnect

SCTP solves most of the above, but nobody is using it

The Consequences of Just Good Enough

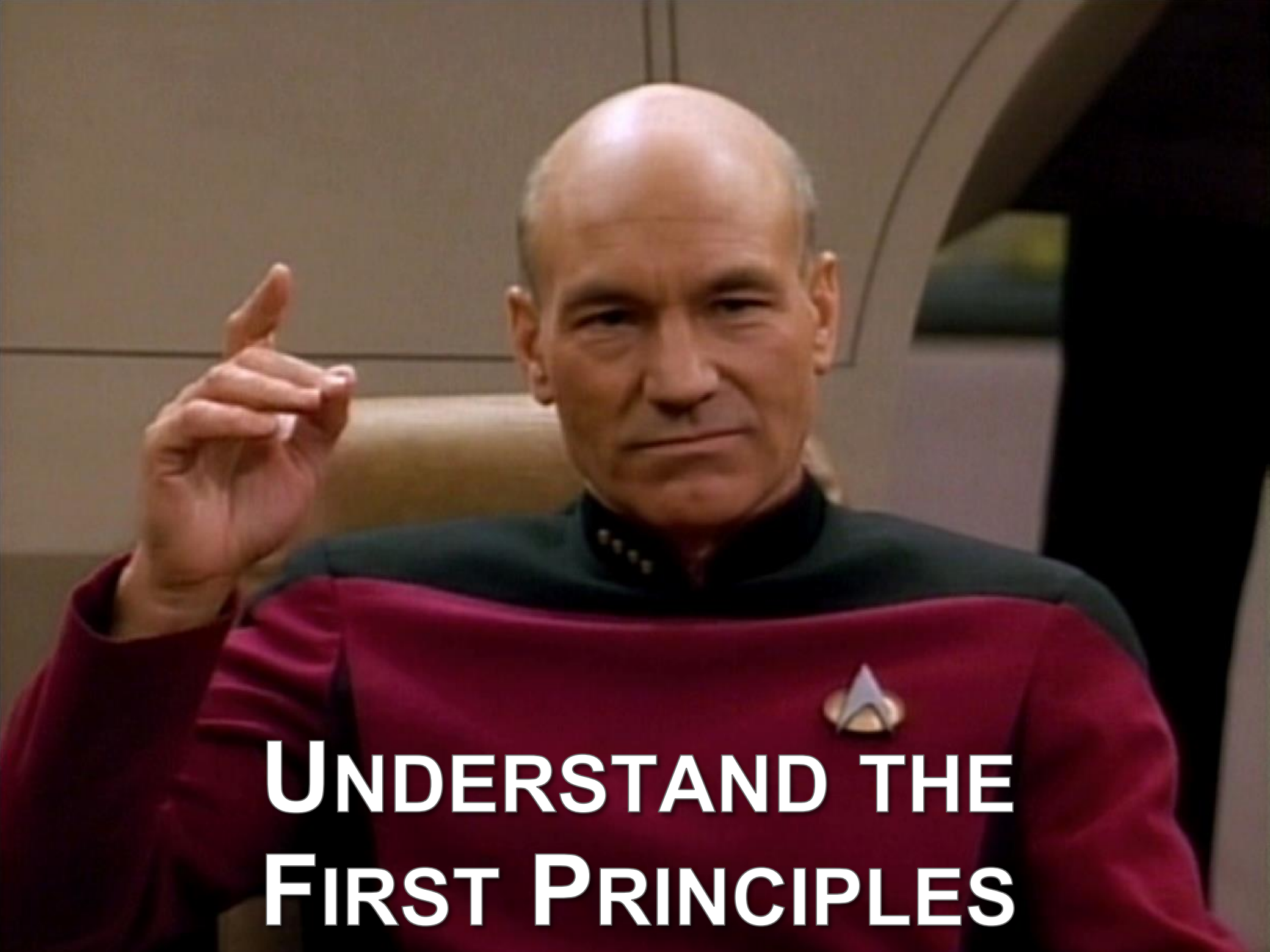
With sufficient thrust, pigs fly just fine.
However, this is not necessarily a good idea.
It is hard to be sure where they are going to
land, and it could be dangerous sitting under
them as they fly overhead.

RFC 1925, Section 2.3

A close-up, high-contrast image of Morpheus from the movie The Matrix. He is bald, has a serious expression, and is wearing dark sunglasses. The reflection in the sunglasses shows two other characters from the movie. The background is a blurred greenish-blue.

**WHAT IF YOU
WANT TO KNOW**

**HOW NETWORKING
REALLY WORKS?**



**UNDERSTAND THE
FIRST PRINCIPLES**

ONE DOES NOT SIMPLY

IGNORE THE HISTORY

The Perils of Ignoring the History

Every old idea will be proposed again with a different name and a different presentation, regardless of whether it works.

RFC 1925, Section 2.11

A photograph of a woman with long dark hair and sunglasses perched on her head, looking towards a young child. The child, who has dark skin and is wearing a striped shirt, stands with their arms crossed and a serious expression. The background is a rustic outdoor setting with a bamboo fence and dirt ground.

ALWAYS WONDER...

AND ASK "WHY?"

It Will Take Time

No matter how hard you try, you can't make a baby in much less than 9 months.

RFC 1925, Section 2.2a

Practice Makes Perfect

Some things in networking can never be fully understood by someone who neither builds commercial networking equipment nor runs an operational network.

RFC 1925, Section 2.4

MUCH TO LEARN...

...YOU HAVE

Stay in Touch

Web: ipSpace.net
Blog: blog.ipSpace.net
Email: ip@ipSpace.net
Twitter: [@ioshints](https://twitter.com/ioshints)



SDN: ipSpace.net/SDN
Webinars: ipSpace.net/Webinars
Consulting: ipSpace.net/Consulting