https://hao-ai-lab.github.io/dsc204a-w24/

DSC 204A: Scalable Data Systems
Winter 2024

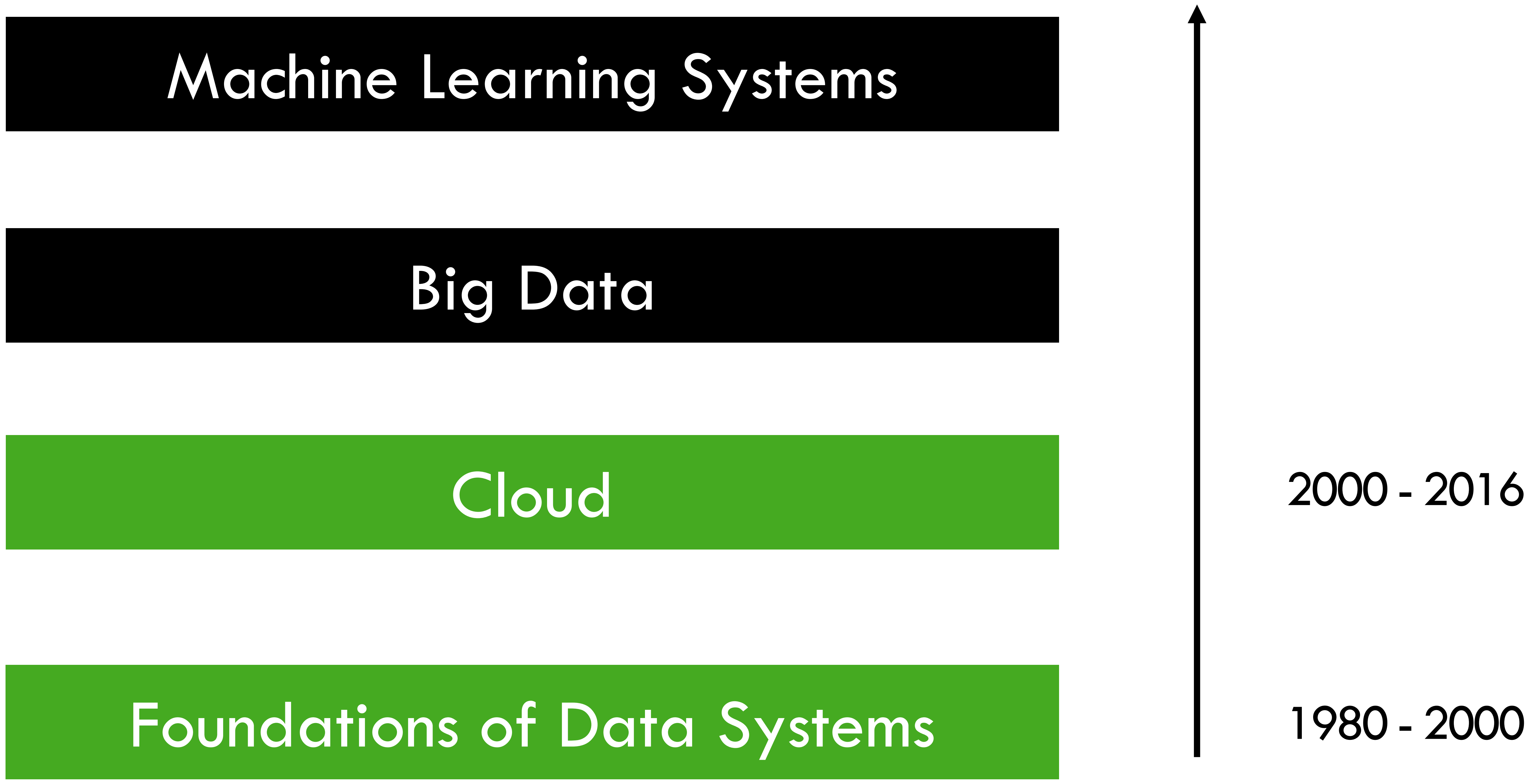
Machine Learning Systems

Big Data

Cloud

Foundations of Data Systems

Where We Are


Machine Learning Systems
Big Data
Cloud
2000 - 2016
Foundations of Data Systems
1980 - 2000

Logistics

- PA1 is up! (Due in 2 weeks)
- TAs has offered one recitation this week to help you learn how to program Ray
  - Tuesday 4 – 5 pm
- TAs have also put up a poll about future recitations.
  - http://tinyurl.com/dsc204a-recitation-poll
- Make sure to provide your preferences

Review Questions

1. Briefly explain 1 pro and 1 con of On-Demand vs Spot instances on AWS.
2. Briefly explain 2 pros and 2 cons of cloud vs on-premise clusters.

Recap: Why Cloud?

Need more compute and storage

Single computer hits physical limits

↓

Distributed Computing

Cloud offers a lot of compute, storage, networking, etc.

Cloud is close to utility computing:

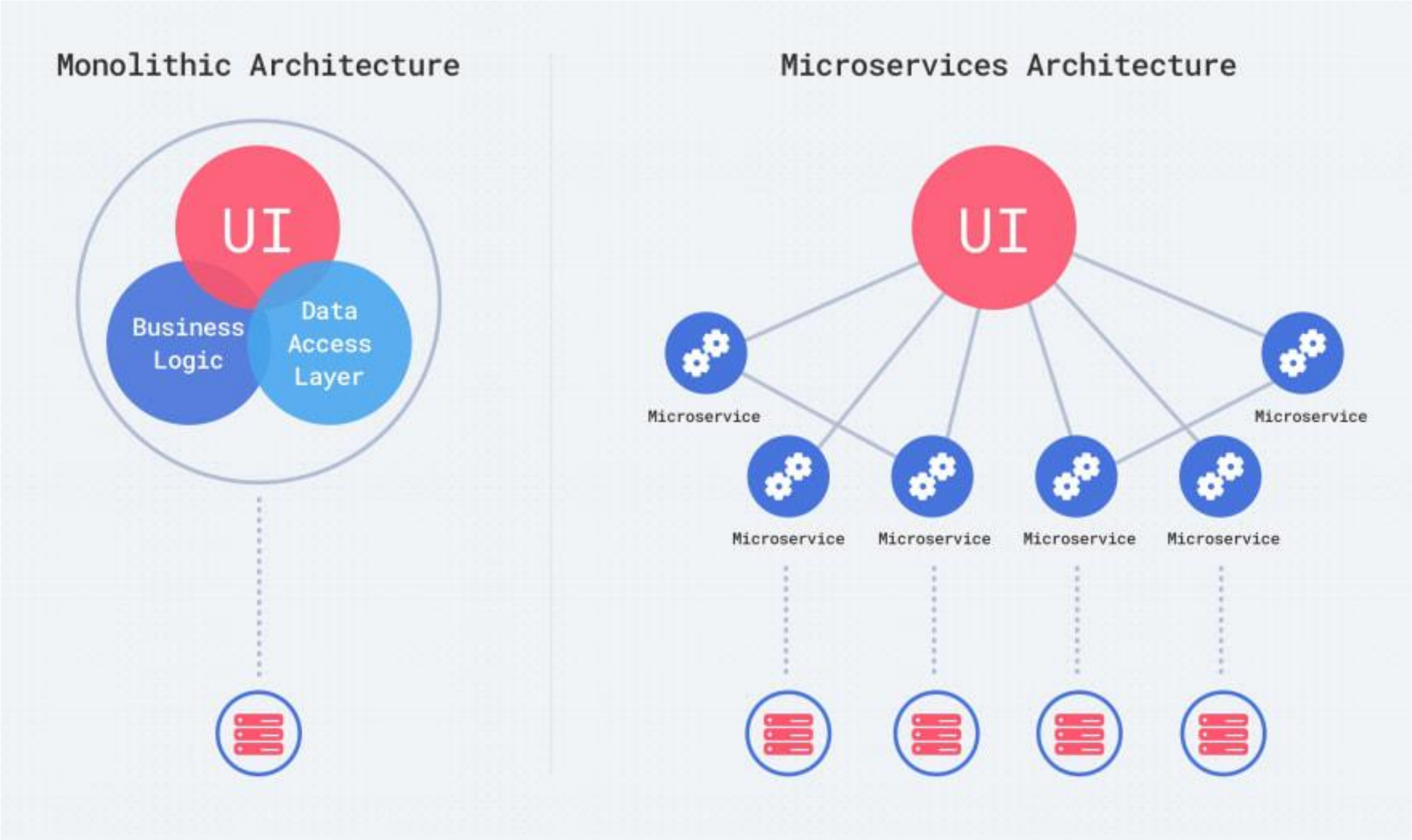
- flexibility and elasticity,
- manageability,
- cost

Open Question after class

Google has pioneered and created many distributed systems and technologies that shape today's cloud computing, but why Amazon (and even Microsoft) wins over Google Cloud (GCP) on Cloud computing market shares?

Instructors' answers: Amazon’s Micro-services vs. Cloud 3.0

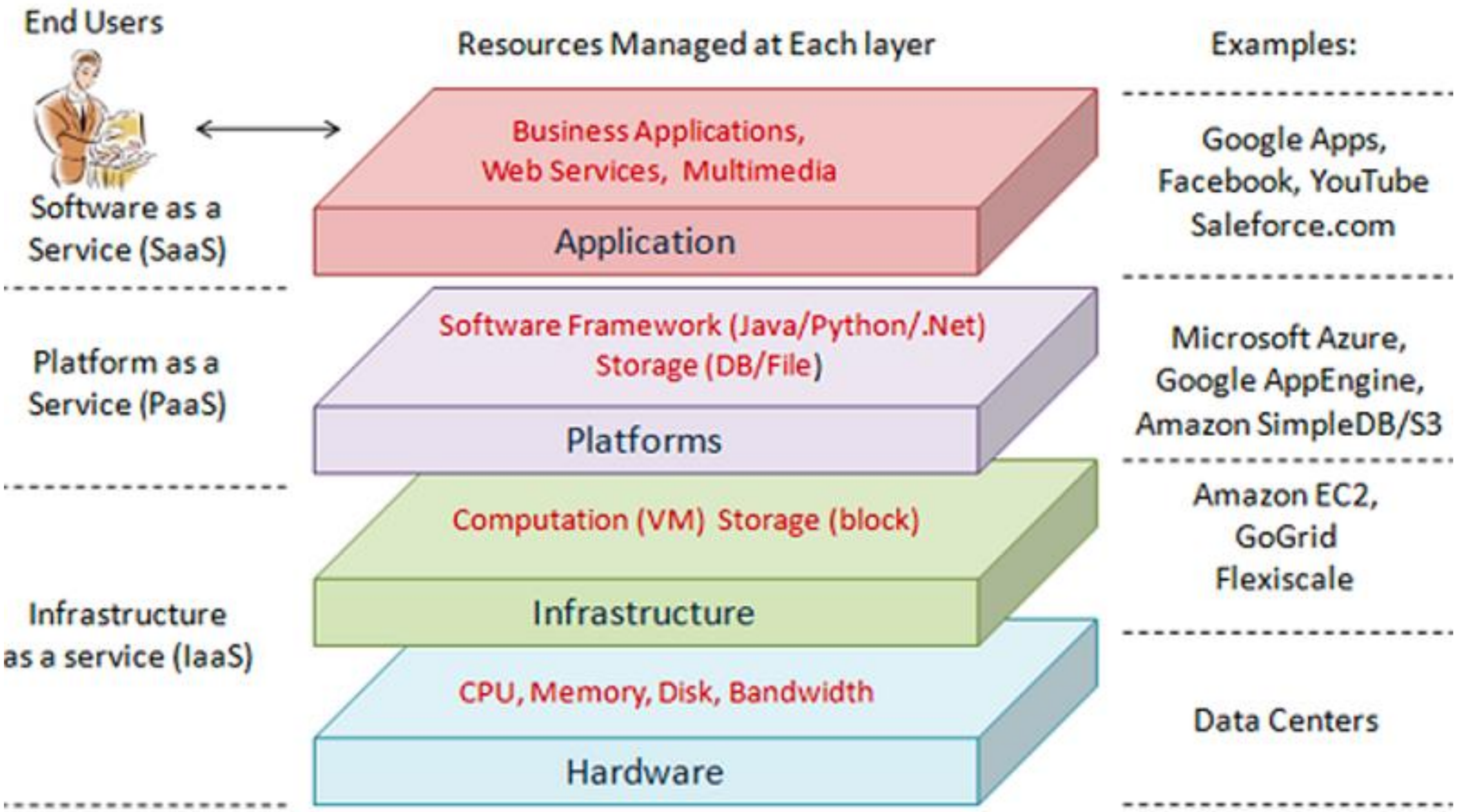
- Microservice architecture correlates well with today’s cloud trend: resources and services are more and more **disaggregated.**



Today's topic

- Why cloud computing?
  - Need-based argument
  - Utility-based argument
- High-level Introduction of Cloud Computing:
  - Cloud computing evolution - sharing granularity
  - Advantages of Cloud computing
  - **Cloud computing layers**

Cloud Computing is a Huge Ecosystem



https://benchpartner.com/layered-model-of-cloud-computing

Examples of AWS Cloud Services

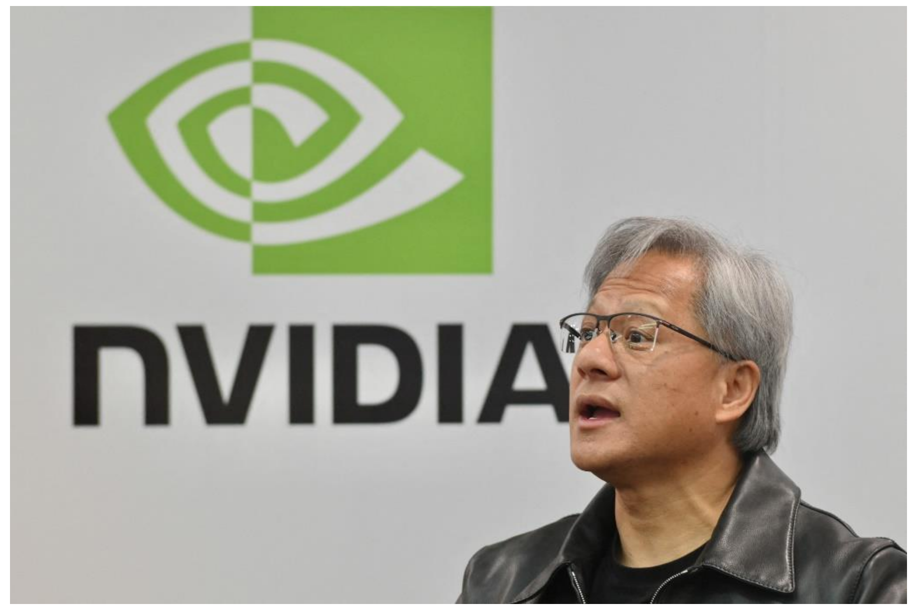
- IaaS: Infra-
  - Compute: EC2, ECS, Fargate, Lambda
  - Storage: S3, EBS, EFS, Glacier
  - Networking: CloudFront, VPC
- PaaS: Platform-
  - Database/Analytics Systems: Aurora, Redshift, Neptune, ElastiCache, DynamoDB, Timestream, EMR, Athena
  - Blockchain: QLDB; IoT: Greengrass
- SaaS: Software-
  - ML/AI: SageMaker, Elastic Inference, Lex, Polly, Translate, Transcribe, Textract, Rekognition, Ground Truth
  - Business Apps: Chime, WorkDocs, WorkMail

However: we are at a transition point, again.

Profit chain

OpenAI | Google DeepMind | Meta

Microsoft | aws | Google Cloud



New Trends in the Deep Learning Era

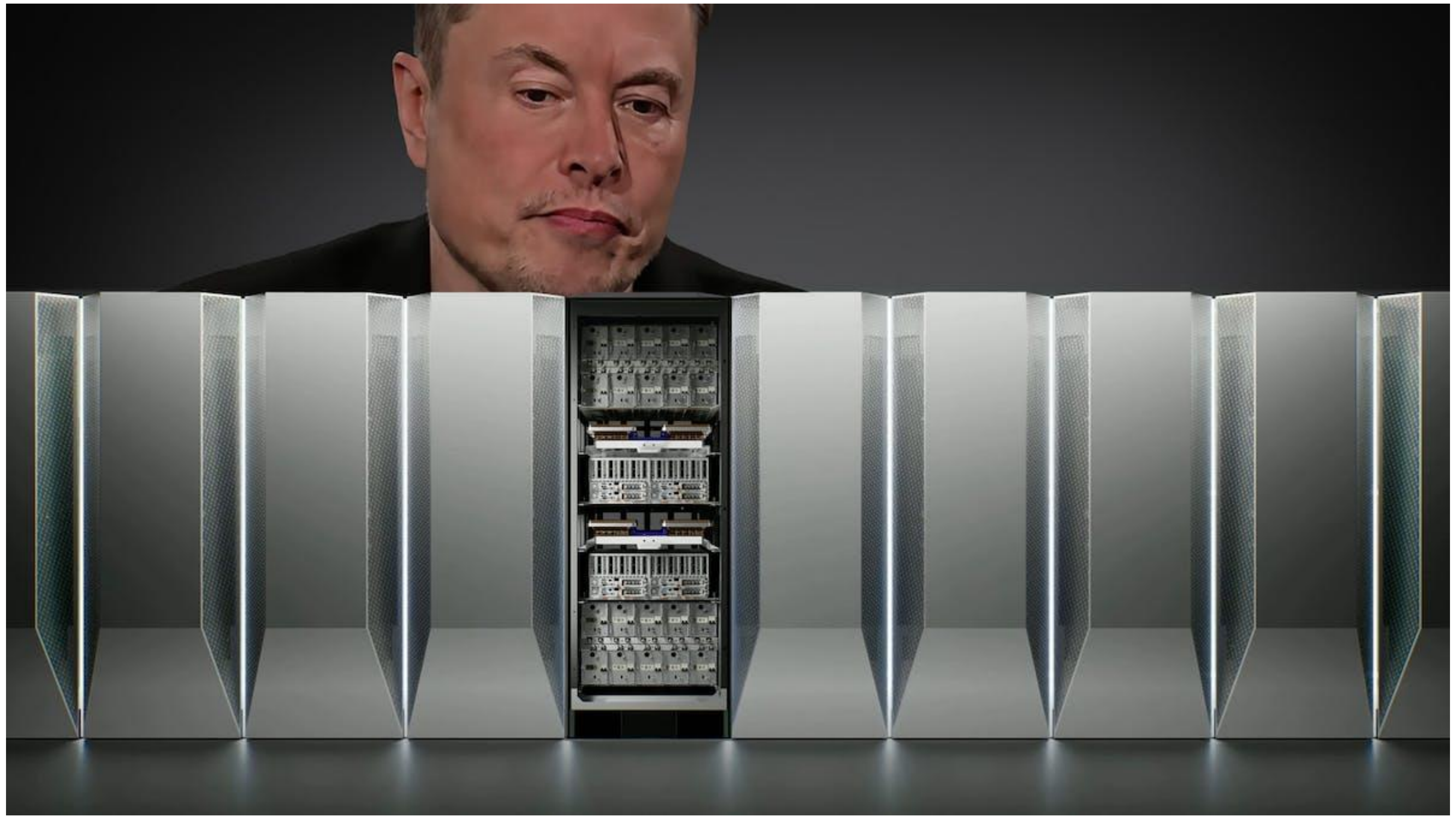
- New trends:
  - Reservation takes precedence than on-demand/spot
  - GPU vertical cloud
  - Community cloud

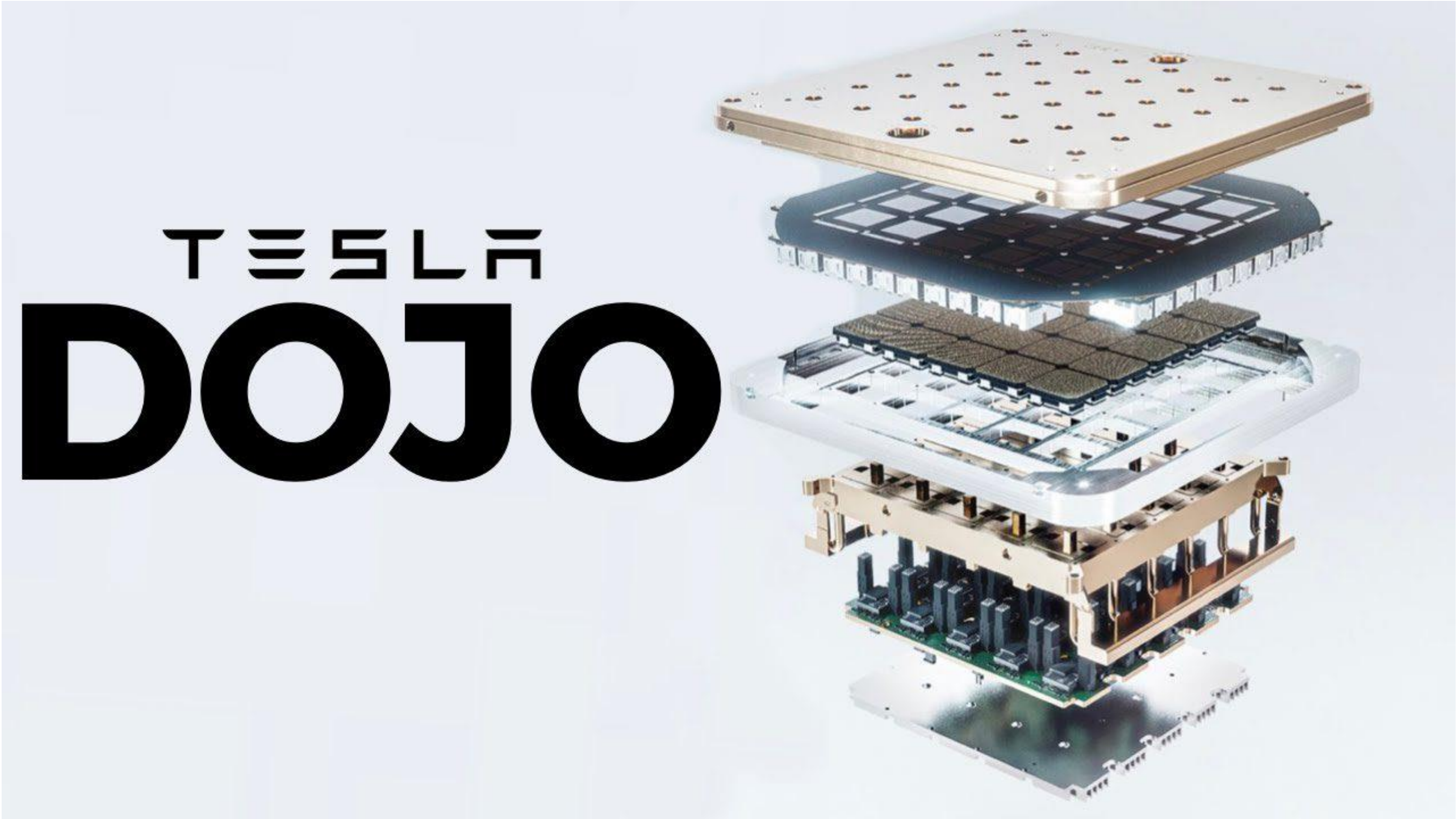
LAST UPDATED: 1 WEEK AGO

Provider	A100 40GB	A100 80GB	H100
Lambda Labs	**Unavailable**	**Unavailable**	**Unavailable**
FluidStack	**Unavailable**	**Unavailable**	**Unavailable**
Runpod	**Unavailable**	$1.59/hour	$3.39/hour

New Trends in the Deep Learning Era

There is a trend of building on-premise super computers again



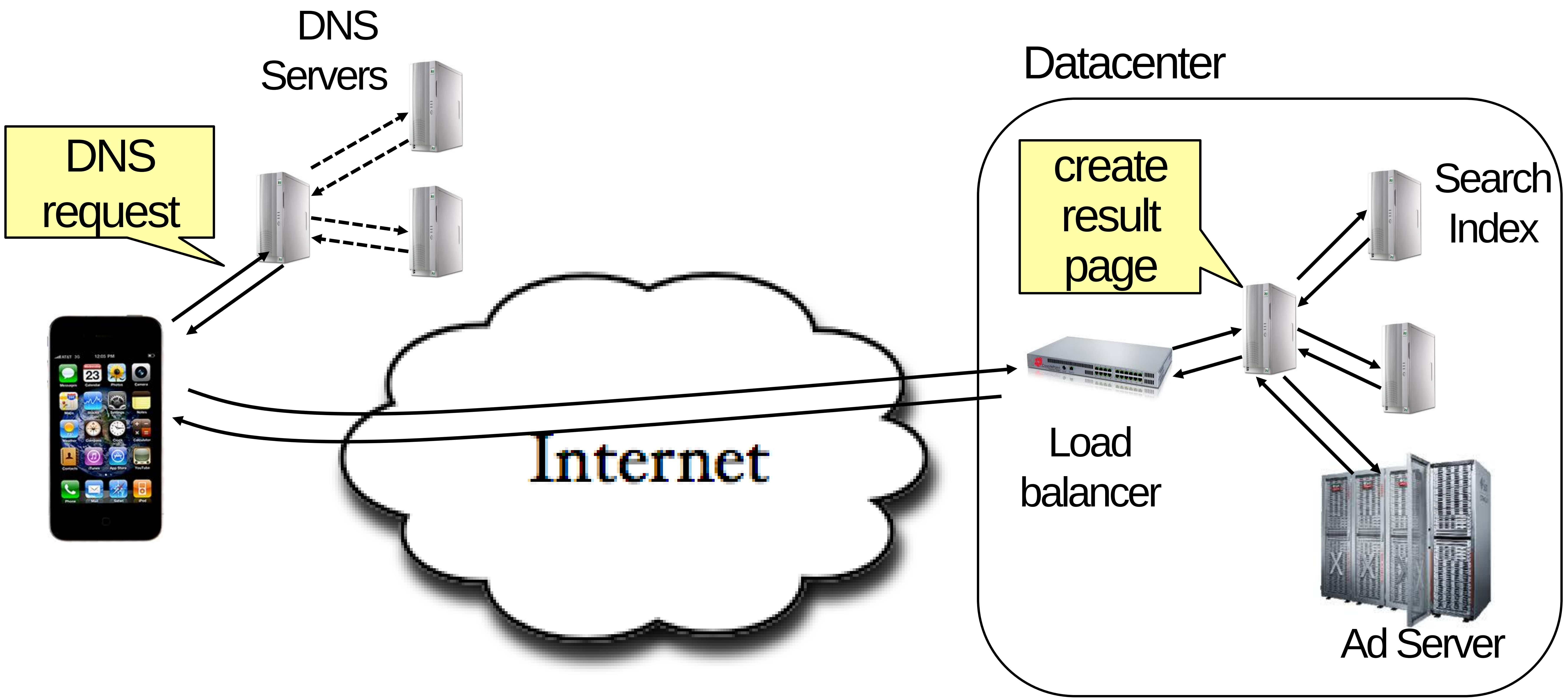


Part 2: Cloud Computing and Distributed Systems

- Intro to Cloud Compute
- **Networking**
- Distributed Storage and file systems
- Distributed Computing
- Parallelism and consistency
- Advanced Topics

Example: What's in a Search Query?

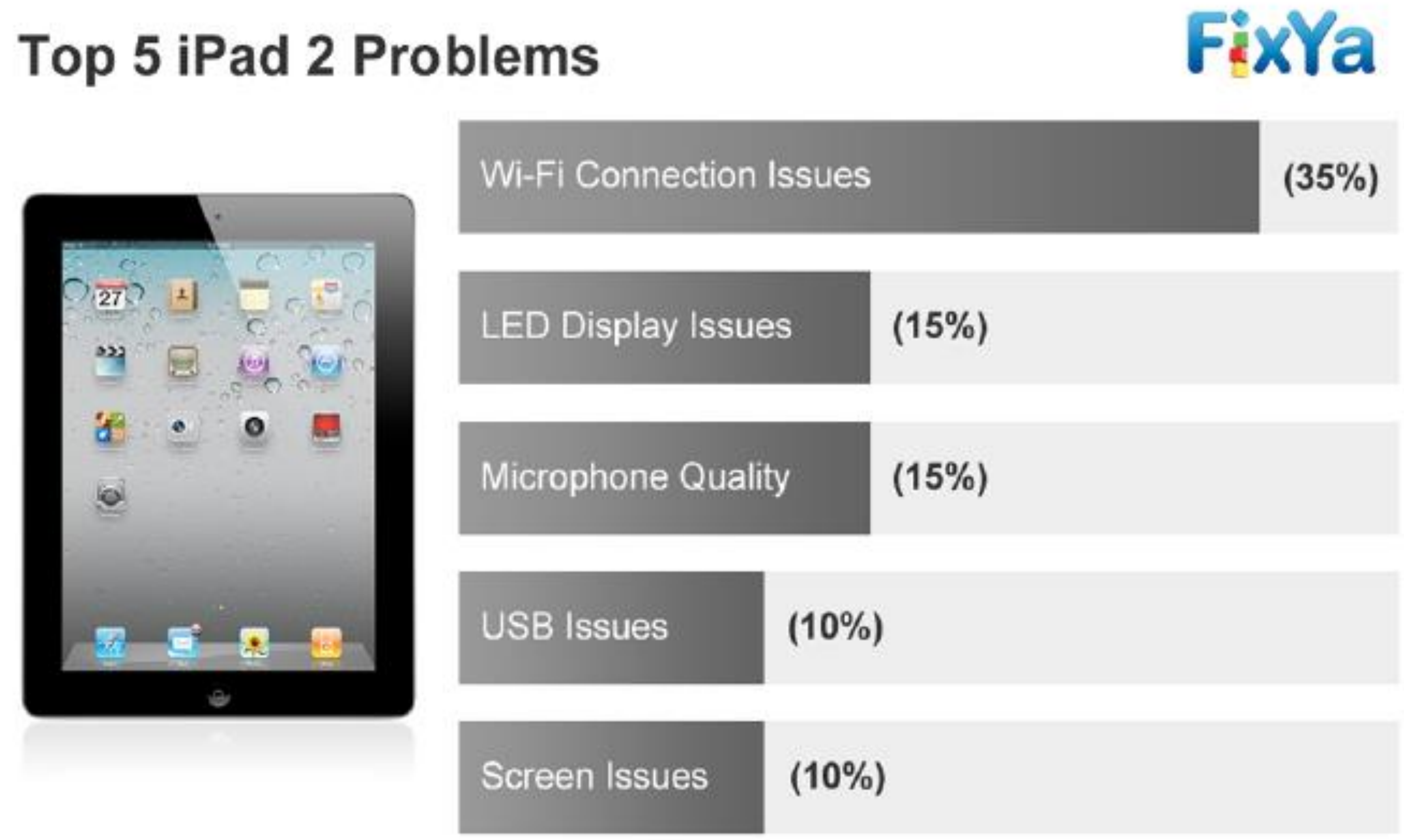
- Complex interaction of multiple components in multiple administrative domains



Why is Networking Imp

- Virtually all apps you use c
  - Many times main functic services, Amazon, Facel

- Thus, connectivity is key se
  - Many times, connectivit



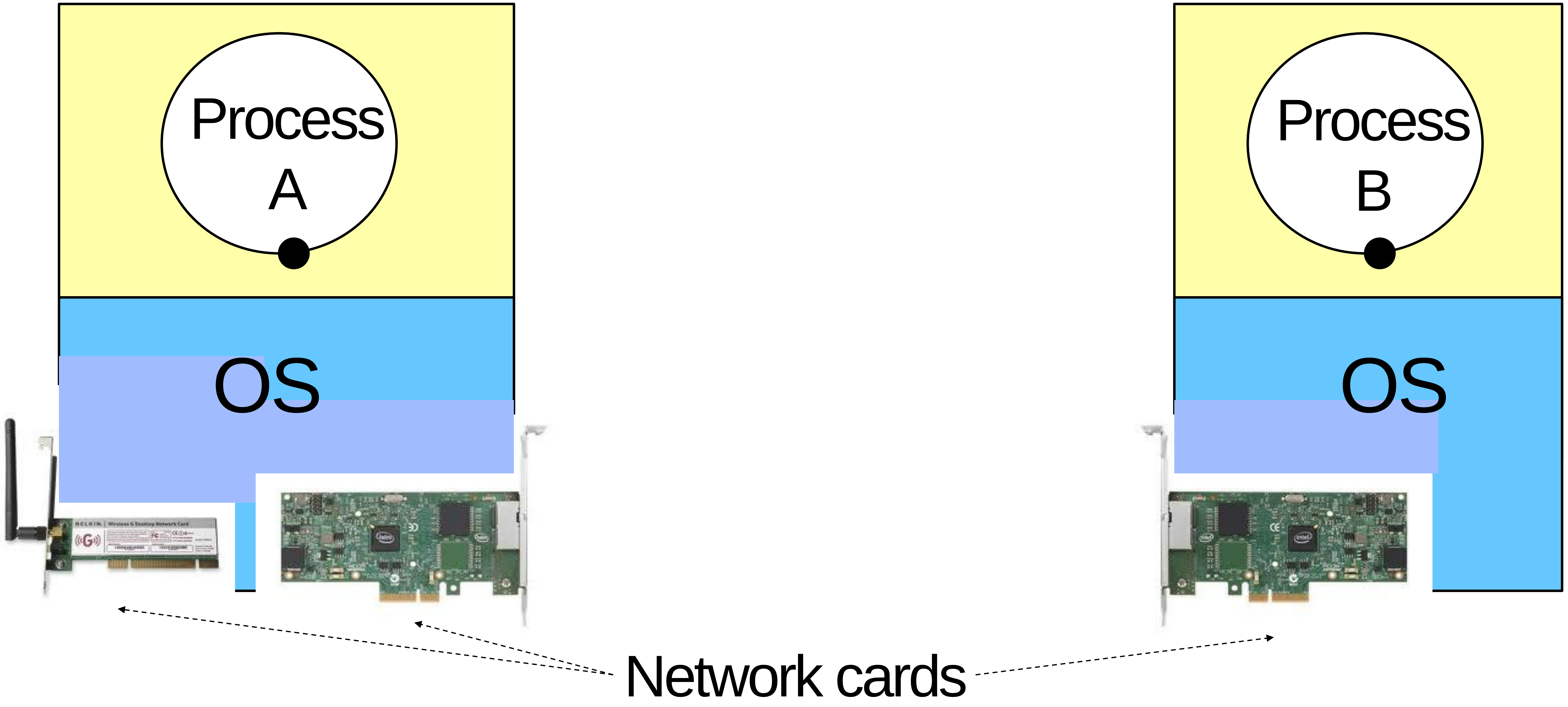
- Some of the hottest opportunities in the OS space:
  - Datacenter networking
  - OSes for Software Defined Networks (SDNs)

Networking

- Network Basics
- Layering and protocols
- Collective Communication

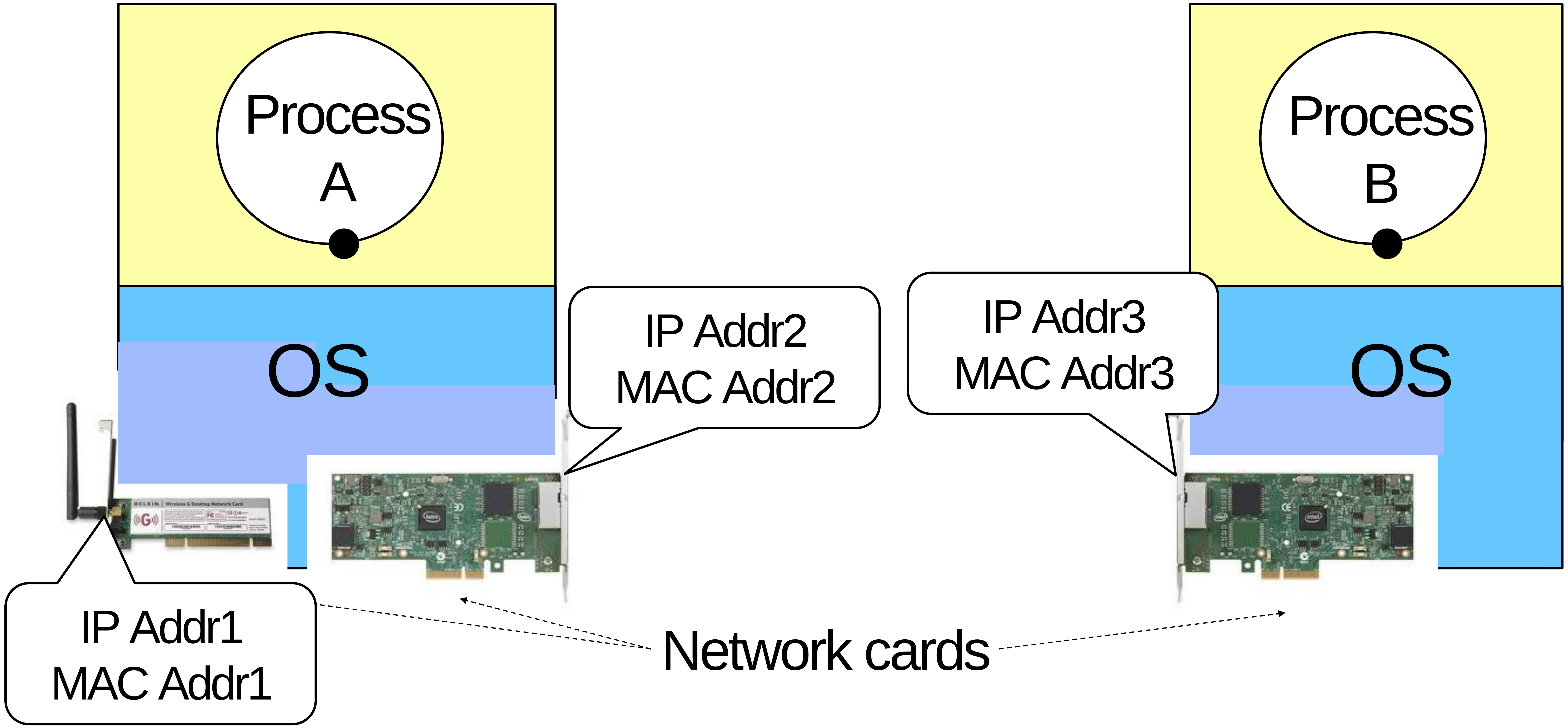
Network Hardware

- **Network (interface) card/controller:** hardware that physically connects a computer to the network



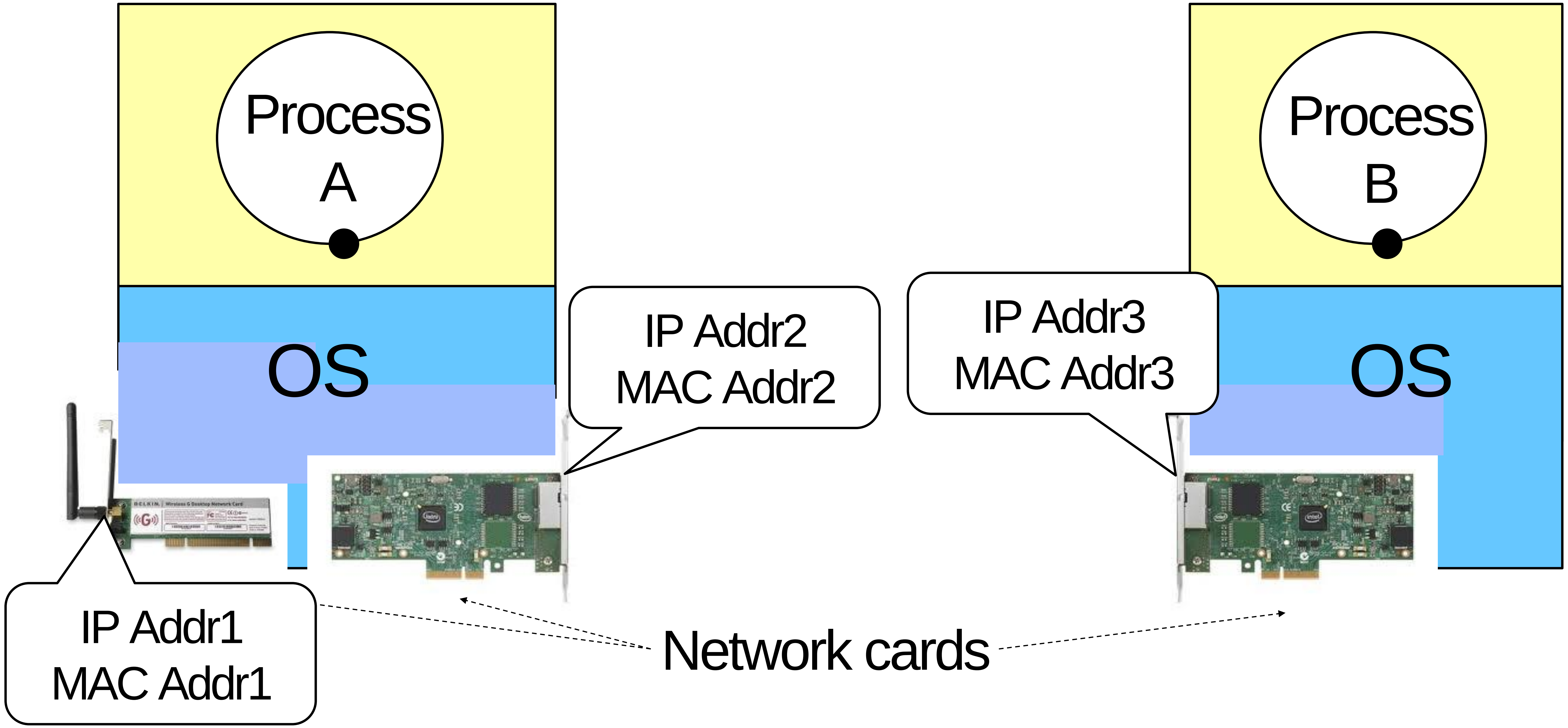
Network Addresses

- Typically, each network card is associated two addresses:
  - Media Access Control (MAC), or physical, address
  - IP (network) address; can be shared by network cards on same host



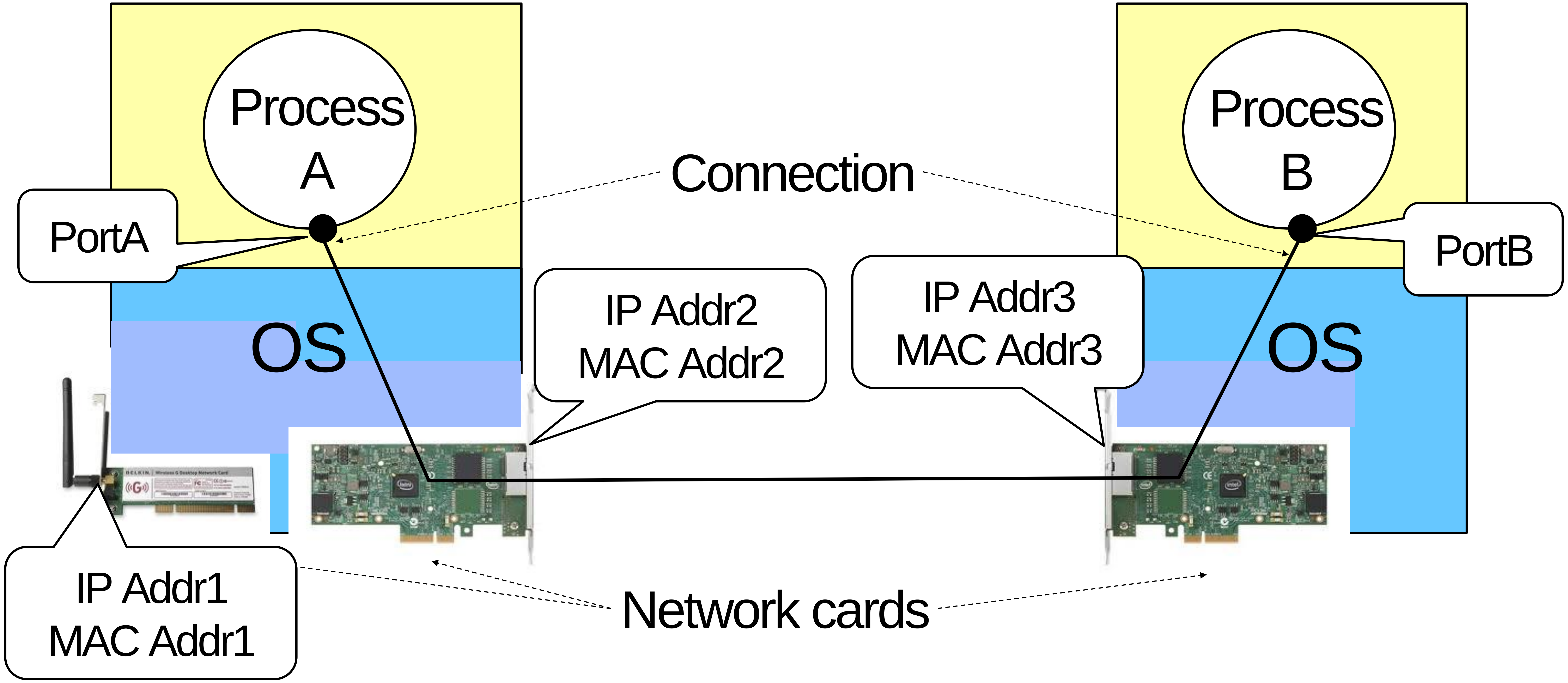
Network Addresses (cont'd)

- **MAC address**: 48-bit unique identifier assigned by card vendor
- **IP Address**: 32-bit (or 128-bit for IPv6) address assigned by network administrator or dynamically when computer connects to network



Network Identifier (cont'd)

- **Connection**: communication channel between two processes
- Each endpoint is identified by a **port number**



Common Port Numbers

Application	Port number
Wake-on-LAN	9
FTP data	20
FTP control	21
SSH	22
Telnet	23
DNS	53
HTTP	80
SNMP	161
…	…

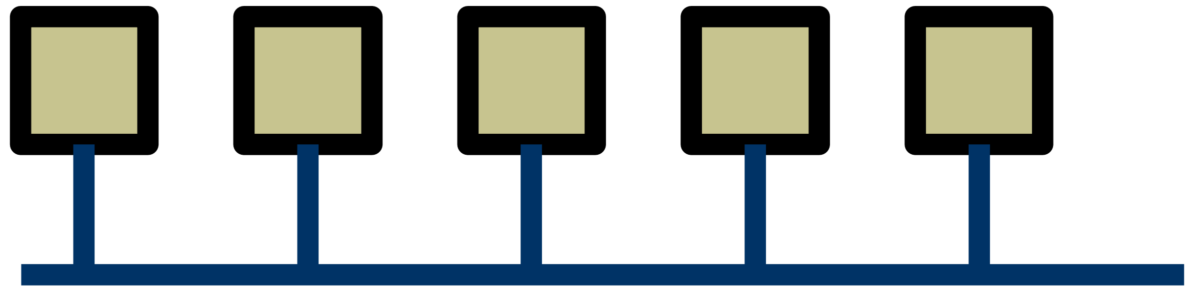
Abstract Network

- Electrical questions
  - Voltage, frequency, …
  - Wired or wireless?

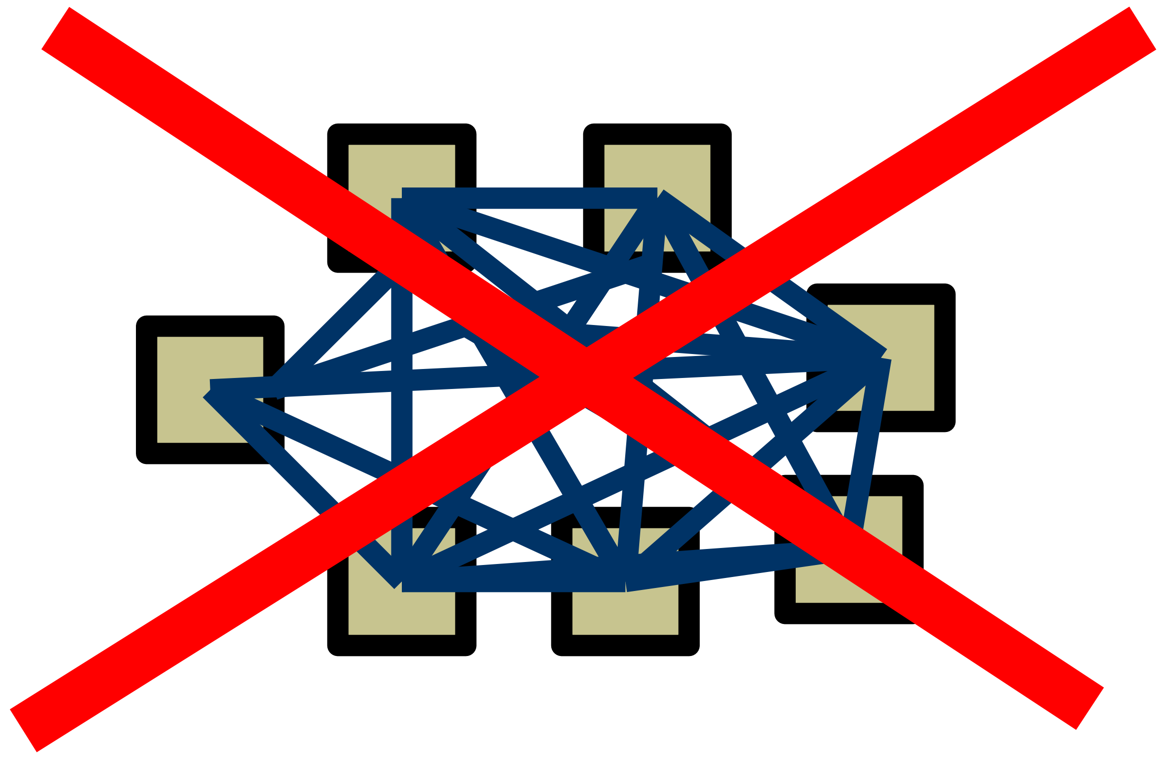
- Link-layer issues: How to send data?
  - When to talk – can either side talk at once?
  - What to say – low-level format?

Basic Building Block: Links

- … But what if we want more hosts?



One wire

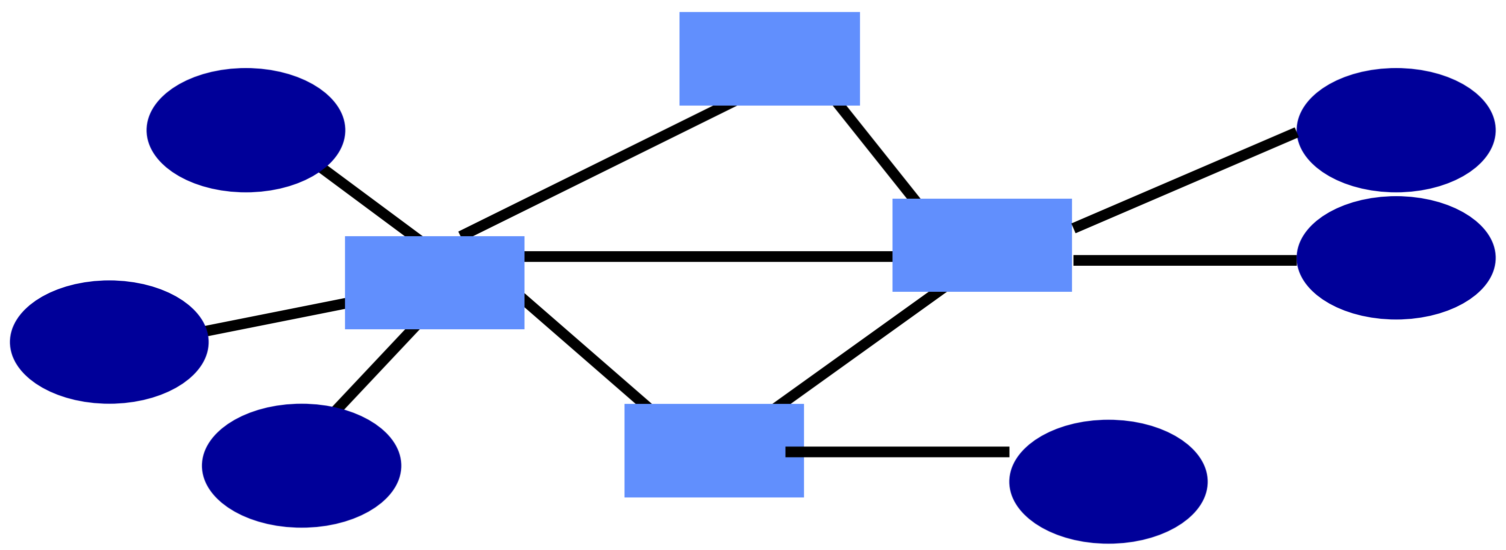


Wires for everybody!

- Scalability?!

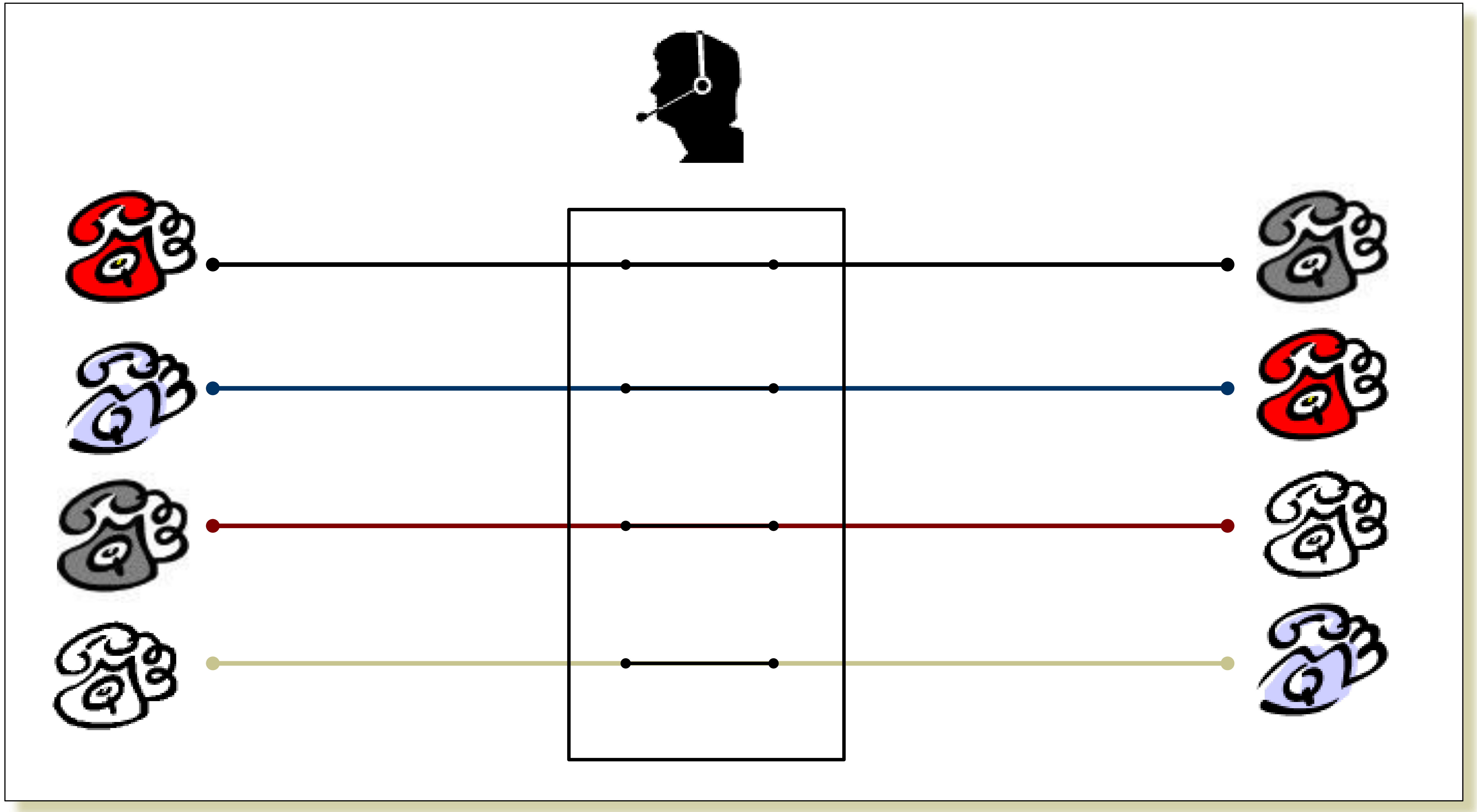
Idea: Multiplexing

- Need to share network resources



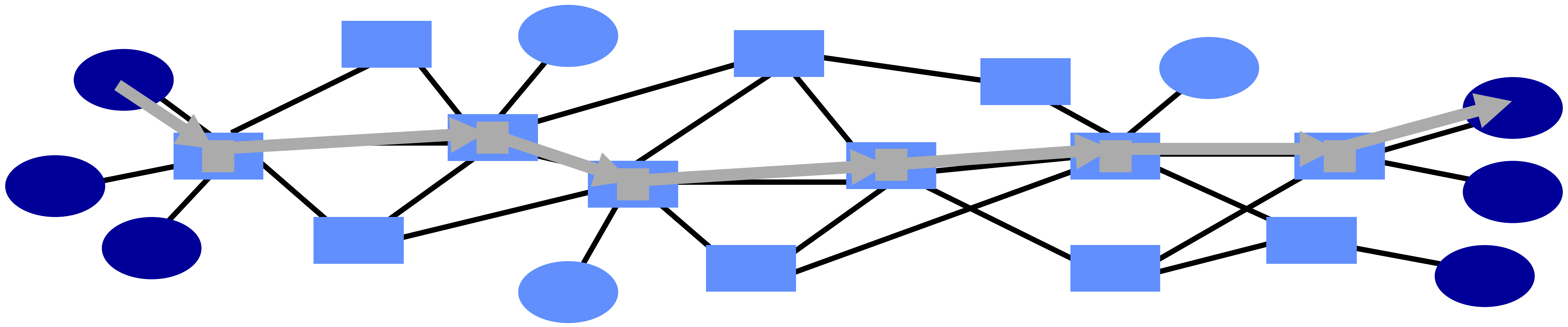
- How? Switched network
  - Party "A" gets resources sometimes
  - Party "B" gets them sometimes
- Interior nodes act as "Switches"
- What mechanisms to share resources?

In the Old Days…Circuit Switching



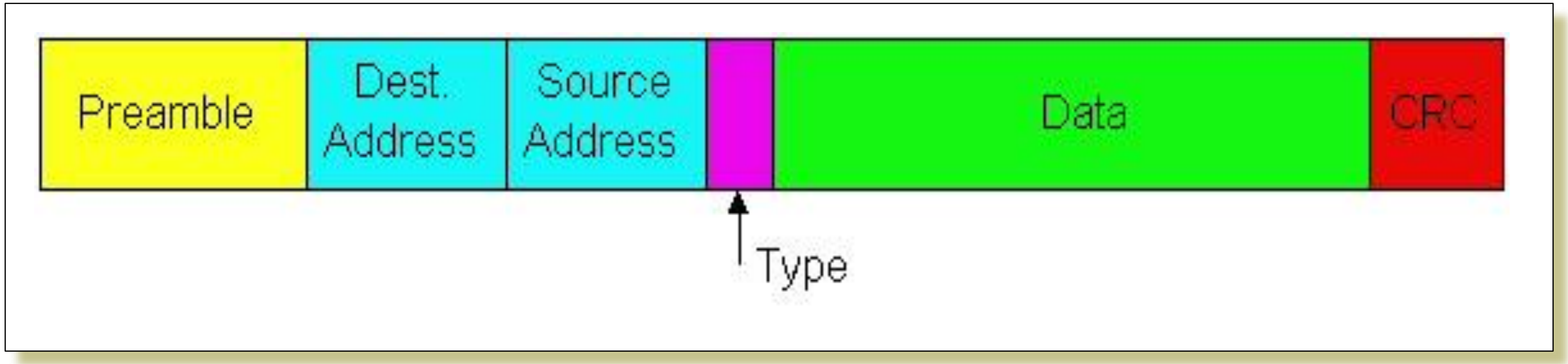
Packet Switching

- Source sends information as self-contained packets that have an address.
  - Source may have to break up single message in multiple packets
- Each packet travels independently to the destination host.
  - Switches use the address in the packet to determine how to forward the packets
  - Store and forward
- Analogy: a letter in surface mail.

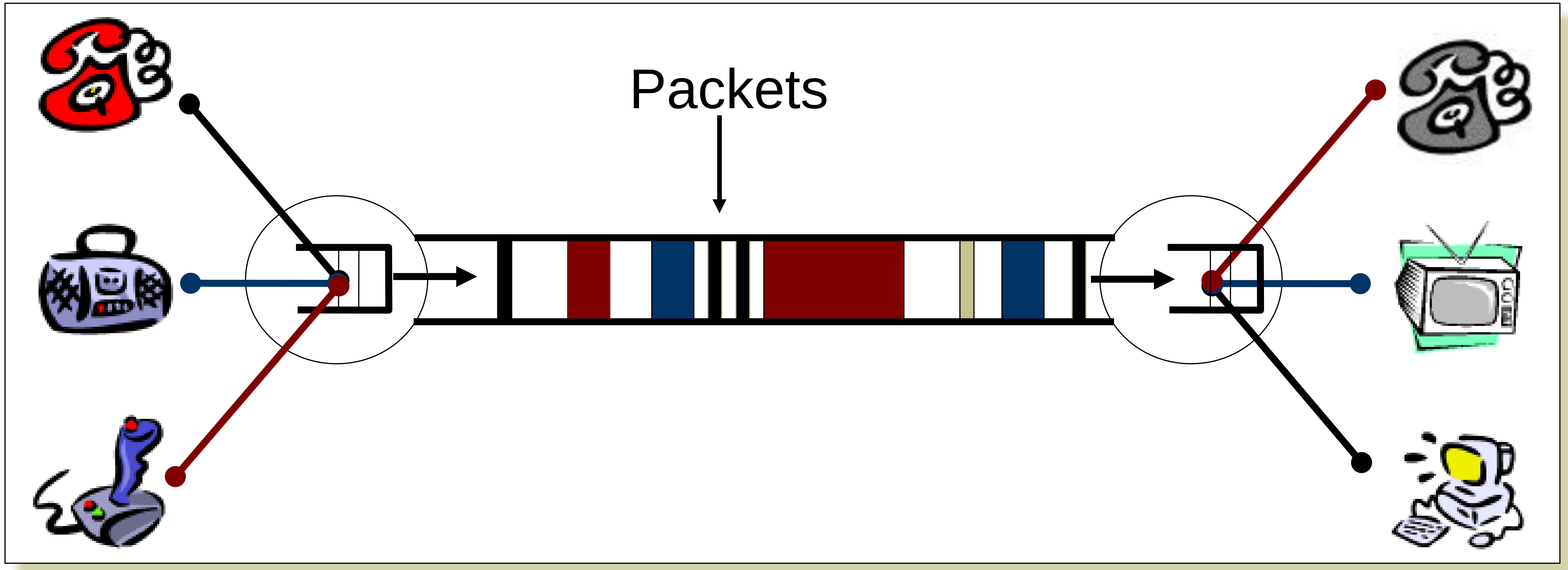


Example: Ethernet Packet

- Sending adapter encapsulates IP datagram (or other network layer protocol packet) in Ethernet frame



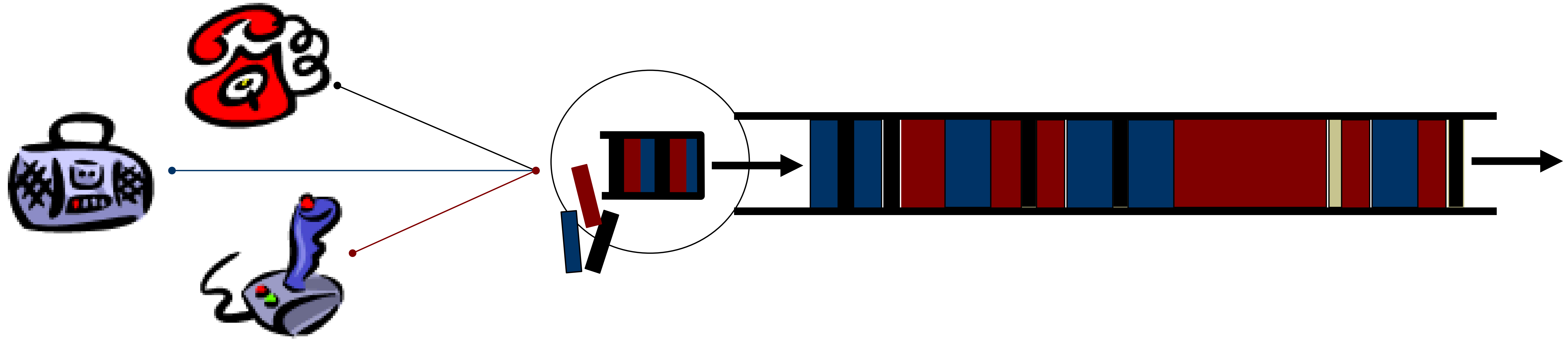
Packet Switching



- Switches arbitrate between inputs
- Can send from *any* input that's ready
  - Links never idle when traffic to send
  - (Efficiency!)

What if Network is Overloaded?

Problem: Network Overload



Solution: Buffering and Congestion Control

- Short bursts: buffer
- What if buffer overflows?
  - Packets dropped
  - Sender adjusts rate until load = resources → "congestion control"

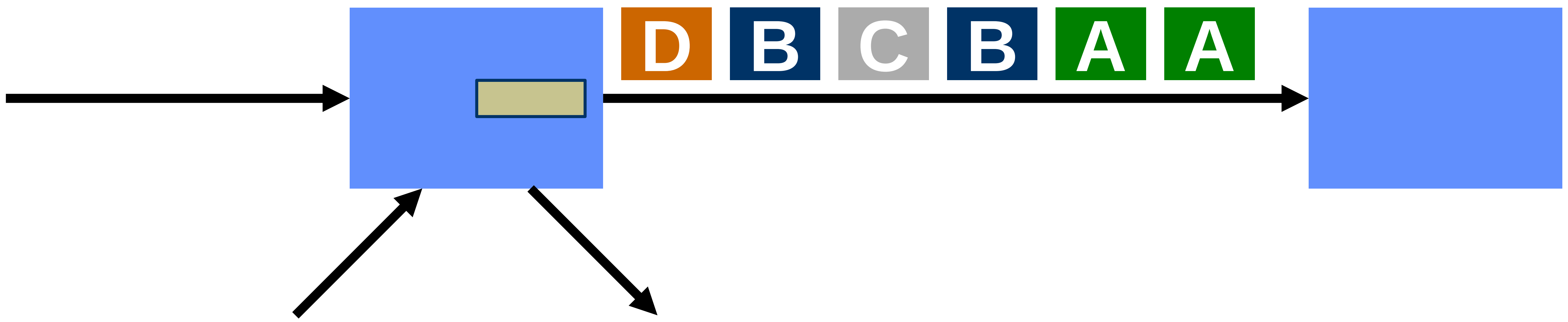
Characterizing Network Communication

- Latency - how long does it take for the first bit to reach destination
- Capacity - how many bits/sec can we push through? (often termed “bandwidth”)
- Jitter - how much variation in latency?
- Loss / Reliability - can the channel drop packets?
- Reordering

Packet Delay

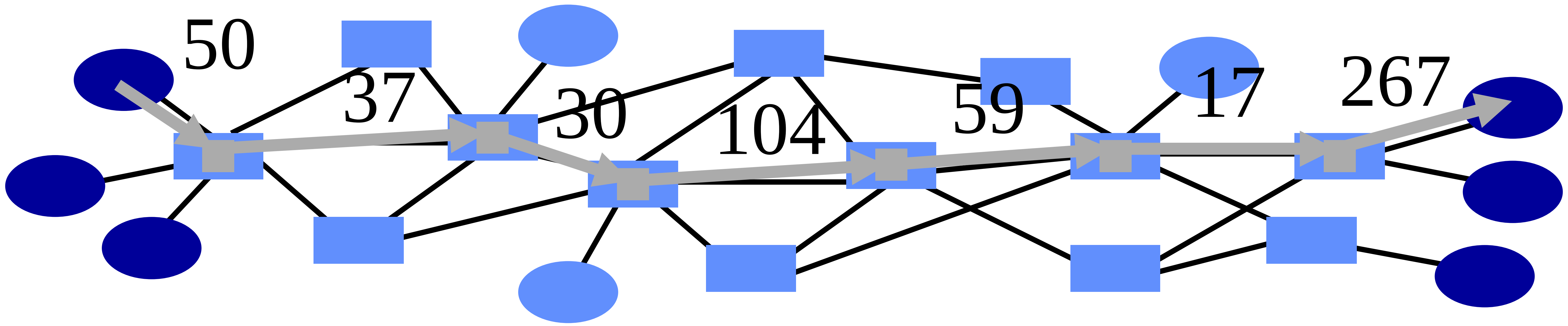
- Sum of a number of different delay components:

- Propagation delay on each link.
  - Proportional to the length of the link
- Transmission delay on each link.
  - Proportional to the packet size and 1/link speed
- Processing delay on each router.
  - Depends on the speed of the router
- Queuing delay on each router.
  - Depends on the traffic load and queue size



Throughput

- When streaming packets, the network works like a pipeline.
  - All links forward different packets in parallel
- Throughput is determined by the slowest stage.
  - Called the bottleneck link
- Does not really matter why the link is slow.
  - Low link bandwidth
  - Many users sharing the link bandwidth



Some simple calculations (mbps/kbps)

- Cross country latency
  - Distance/speed = 5 * 10^6m / 2x10^8m/s = 25 * 10^-3 s = 25ms
  - 50ms RTT
- Link speed (capacity) 100Mbps
- Packet size = 1250 bytes = 10 kbits
  - Packet size on networks usually = 1500 bytes across wide area or 9000 bytes in local area
- 1 packet takes
  - 10k/100M = .1 ms to transmit
  - 25ms to reach there
  - ACKs are small → so 0ms to transmit
  - 25ms to get back
- Effective bandwidth = 10kbits/50.1ms = 200kbits/sec 🐌

https://www.quora.com/Which-is-faster-the-speed-of-electric-current-through-a-copper-wire-or-the-speed-of-l

Some Examples

- How long does it take to send a 100 Kbit file?
  - Assume a perfect world

Latency \ Throughput	100 Kbit/s	1 Mbit/s	100 Mbit/s
500 μsec	1.0005	0.1005	0.0015
10 msec	1.01	0.11	<u>0.011</u>
100 msec	1.1	0.2	<u>0.101</u>

Some Examples

- How long does it take to send a 10 Kbit file?
  - Assume a perfect world

Throughput / Latency	100 Kbit/s	1 Mbit/s	100 Mbit/s
500 μsec	0.1005	0.0105	0.0006
10 msec	0.11	0.02	0.0101
100 msec	0.2	0.11	0.1001