Advanced Computer Networks

# Data Center Network for GPUs (I)

CS740

https://pages.cs.wisc.edu/~mgliu/CS740/F25/index.html

**Ming Liu**
**mgliu@cs.wisc.edu**

# Outline

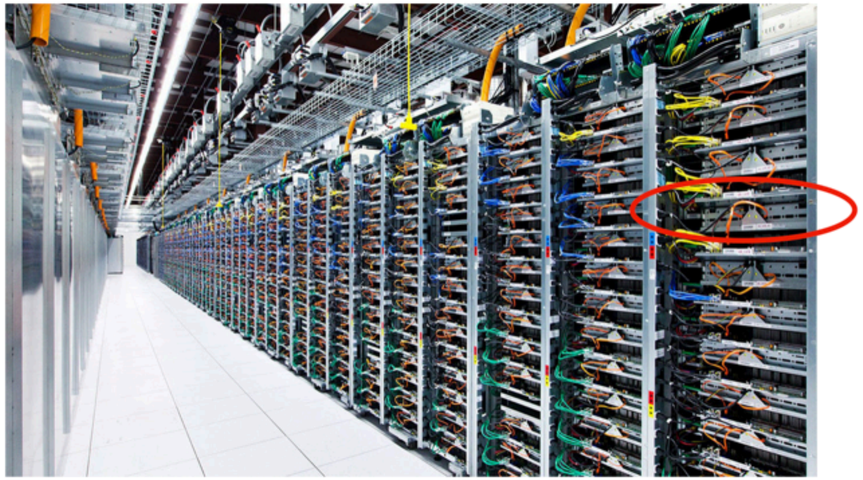
- Last lecture
  - Exam
- Today
  - Data Center Network For GPUs (I)
- Announcements
  - Project Presentation on 12/04/2025 and 12/09/2025

# Building Data Center Networks for GPUs

**L1**

## Data Center Networks

- Data center networks connect ~~servers~~ **GPUs**



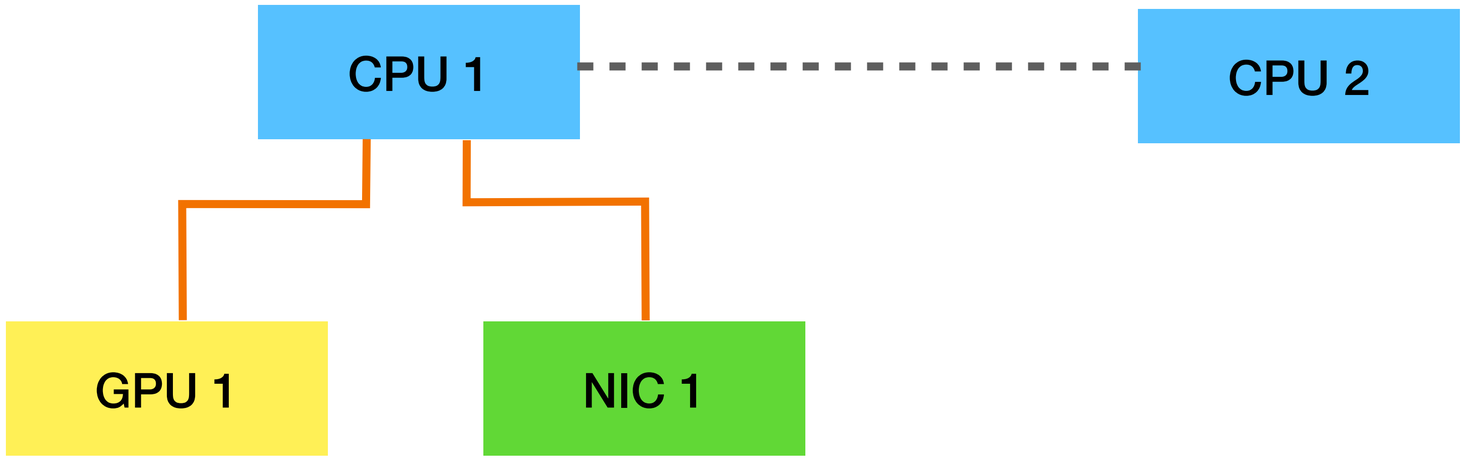
# But why don’t we reuse the existing DCNet?

# GPUs are not standalone devices

- Inside the server: PCIe cards (GPU, FPGA, and ASIC)

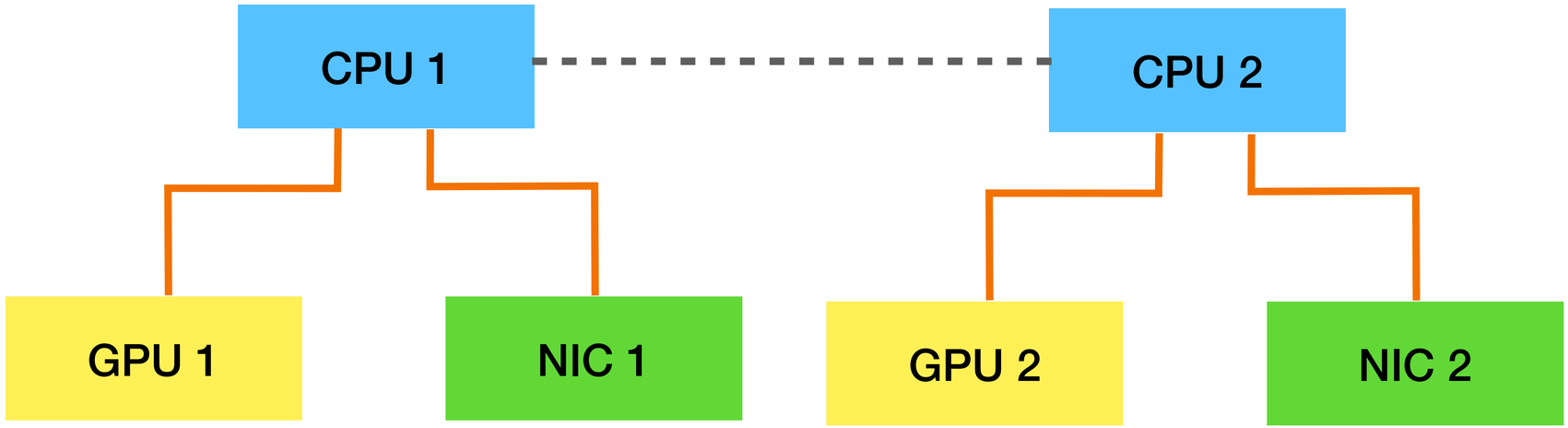
# Server Architecture Block Diagram

- 1-GPU server



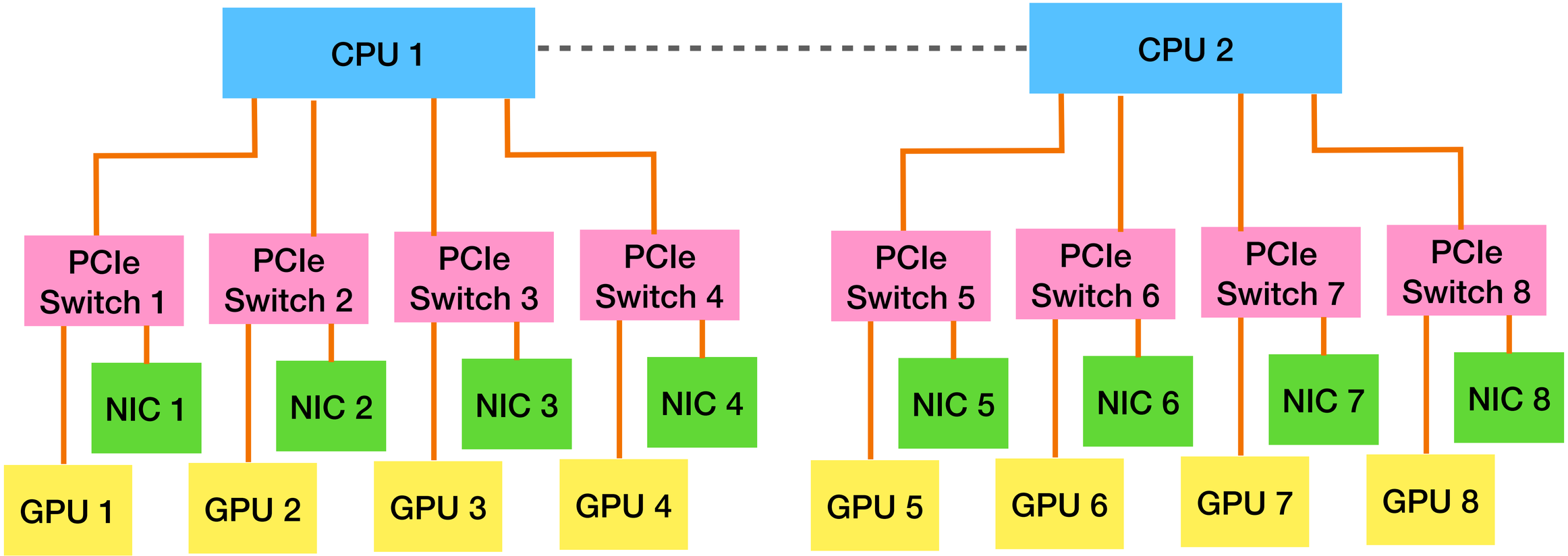
# Server Architecture Block Diagram

- 2-GPU server



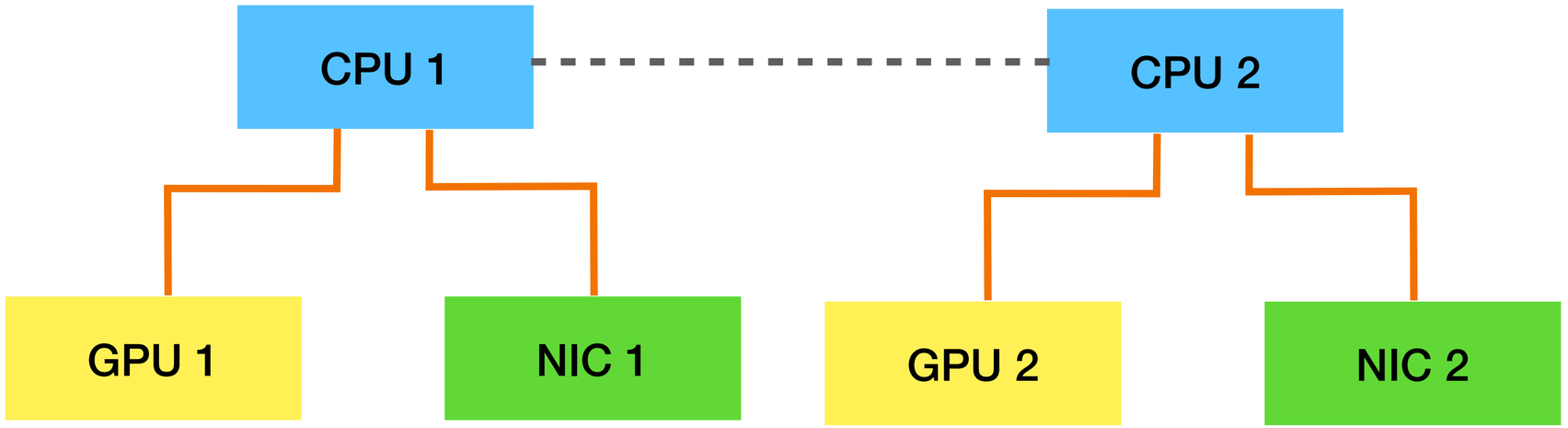
# Server Architecture Block Diagram

- 8-GPU server



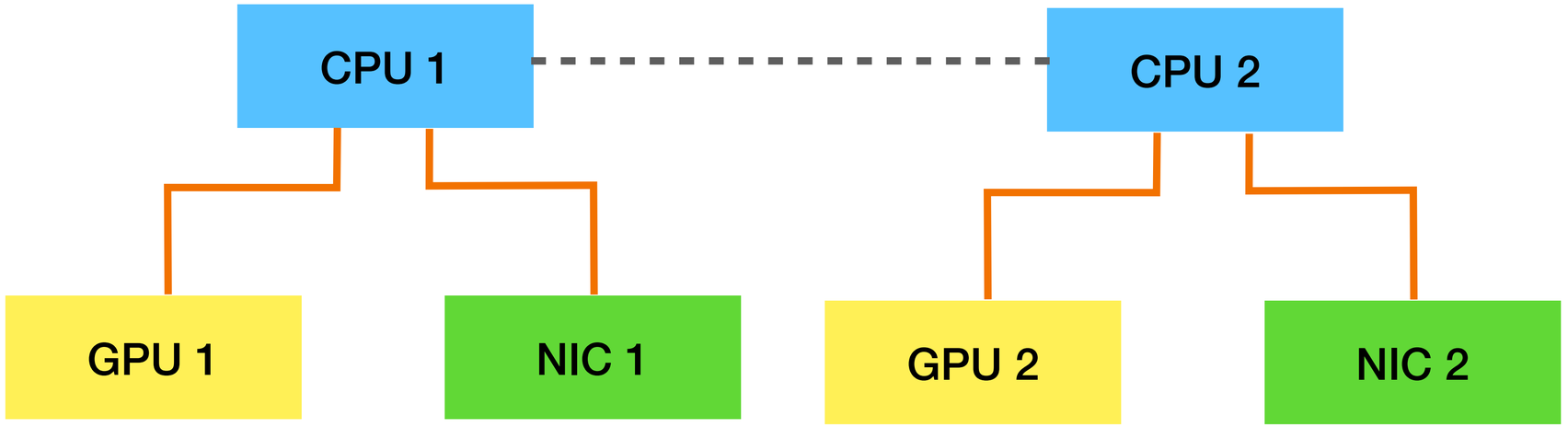
# How does GPU-GPU communication work?

# Case 1: Intra-Server

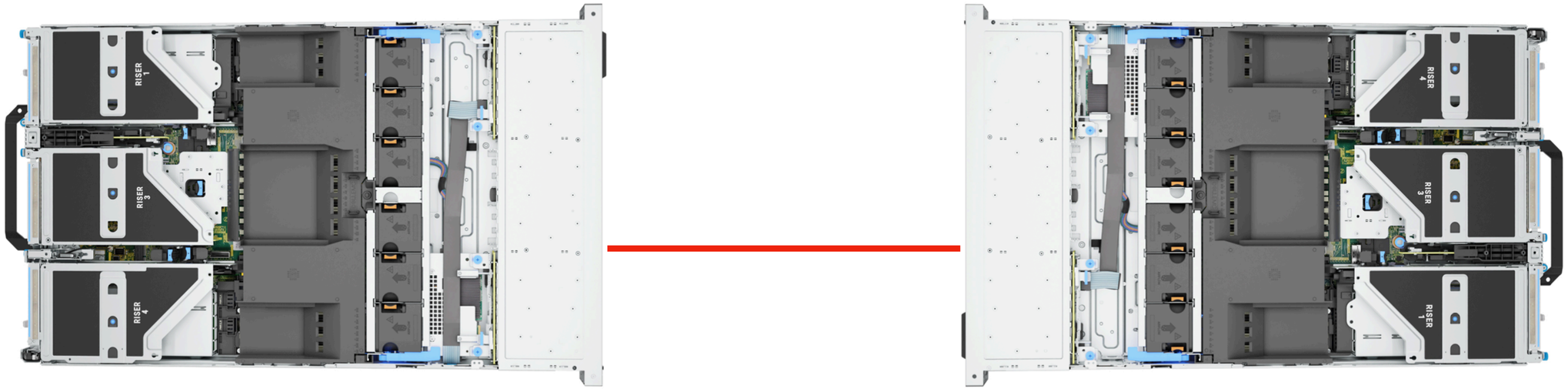


# Case 1: Intra-Server

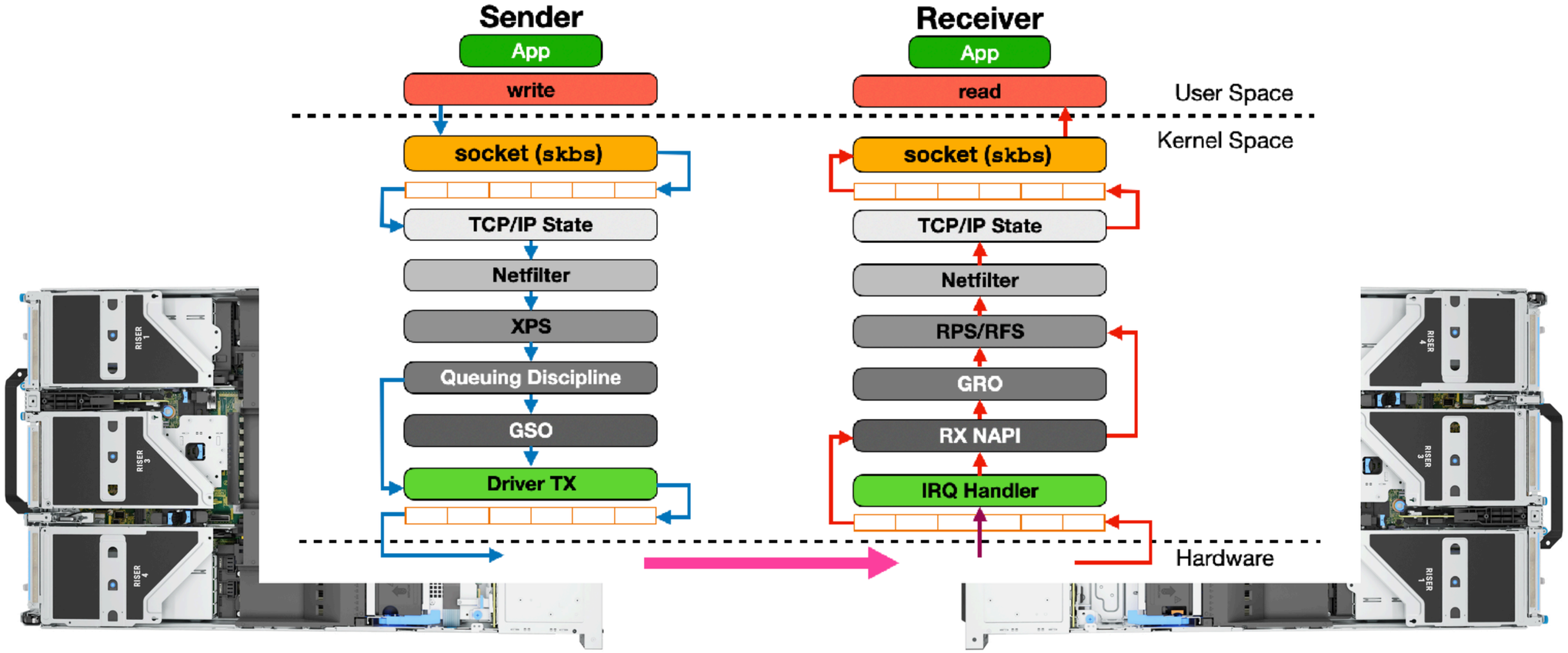
- #1: Host memory bouncing buffer
- #2: PCIe peer-to-peer
- #3: NVLink



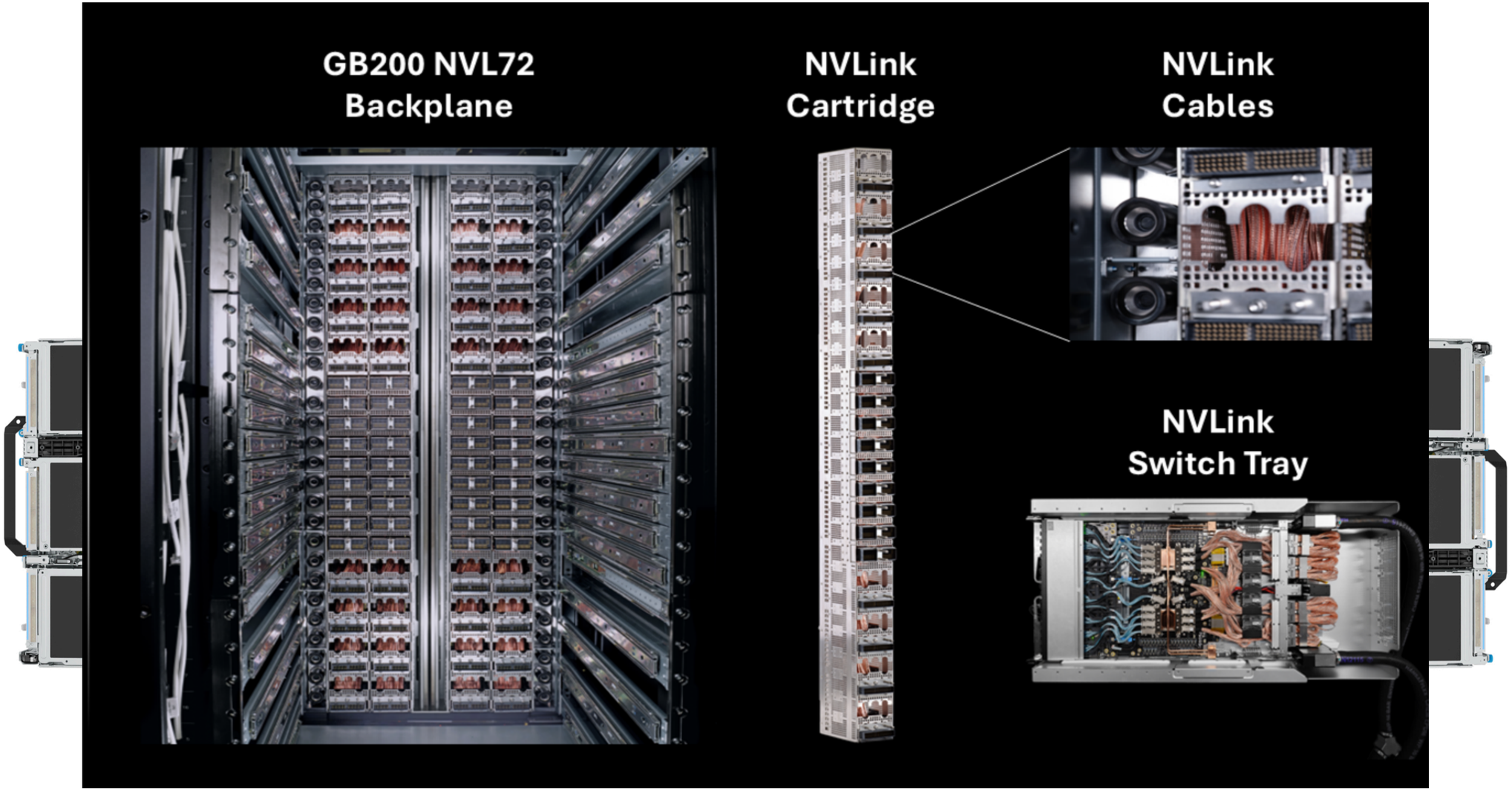
# Case 2: Inter-Server Intra-Rack



# Case 2: Inter-Server Intra-Rack

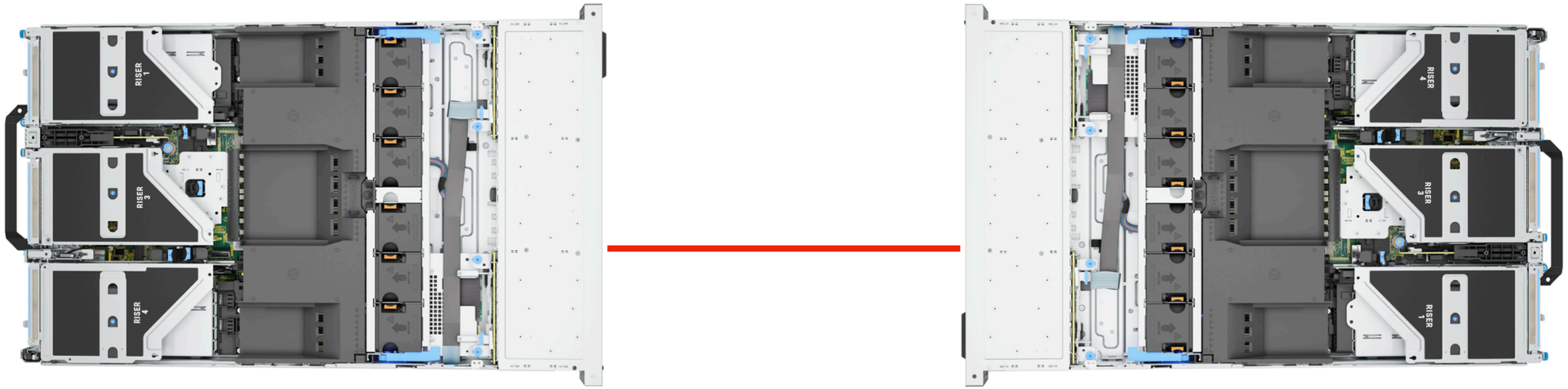


# Case 2: Inter-Server Intra-Rack



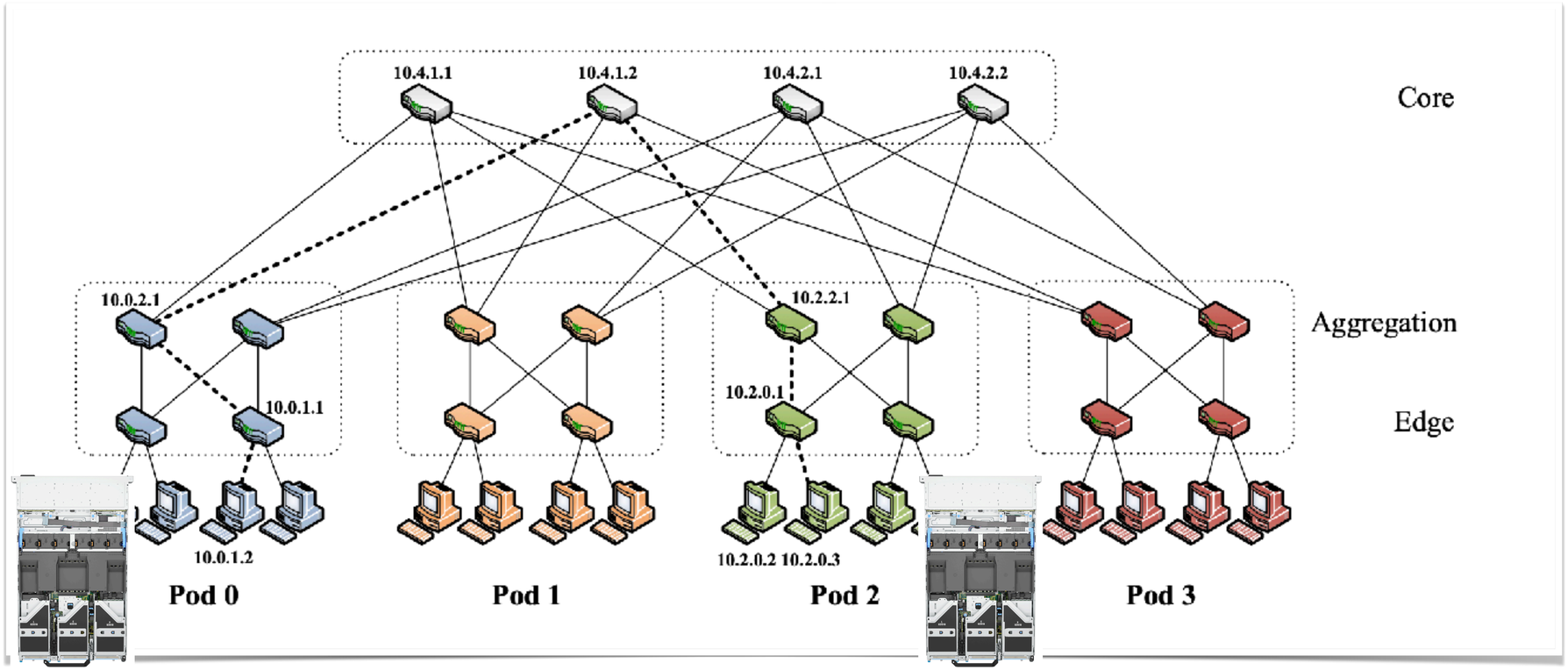
# Case 2: Inter-Server Intra-Rack

- #1: OS networking stack
- #2: Kernel-bypass networking stack (RDMA and Infiniband)
- #3: NVLink



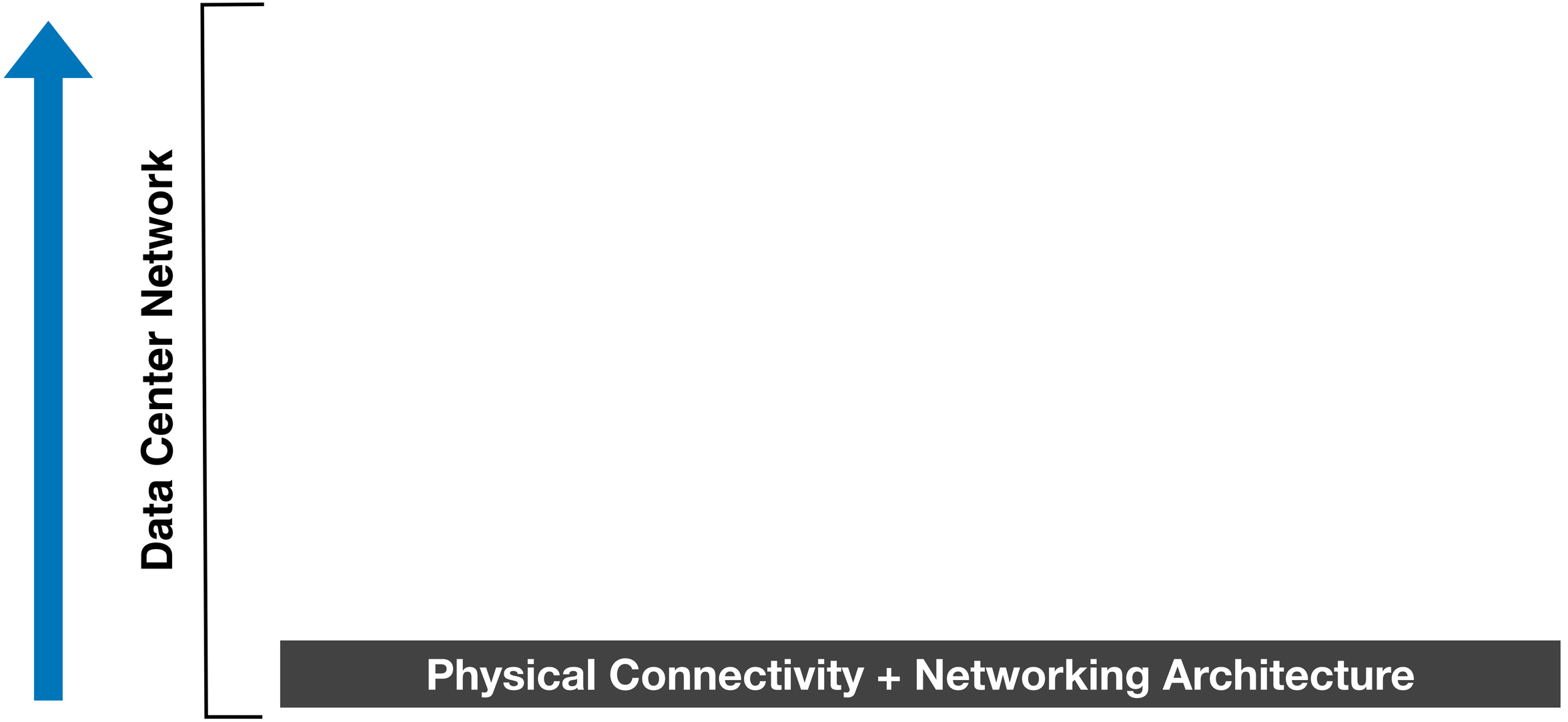
# Case 3: Inter-Server Cross-Rack

- Data Center Networks

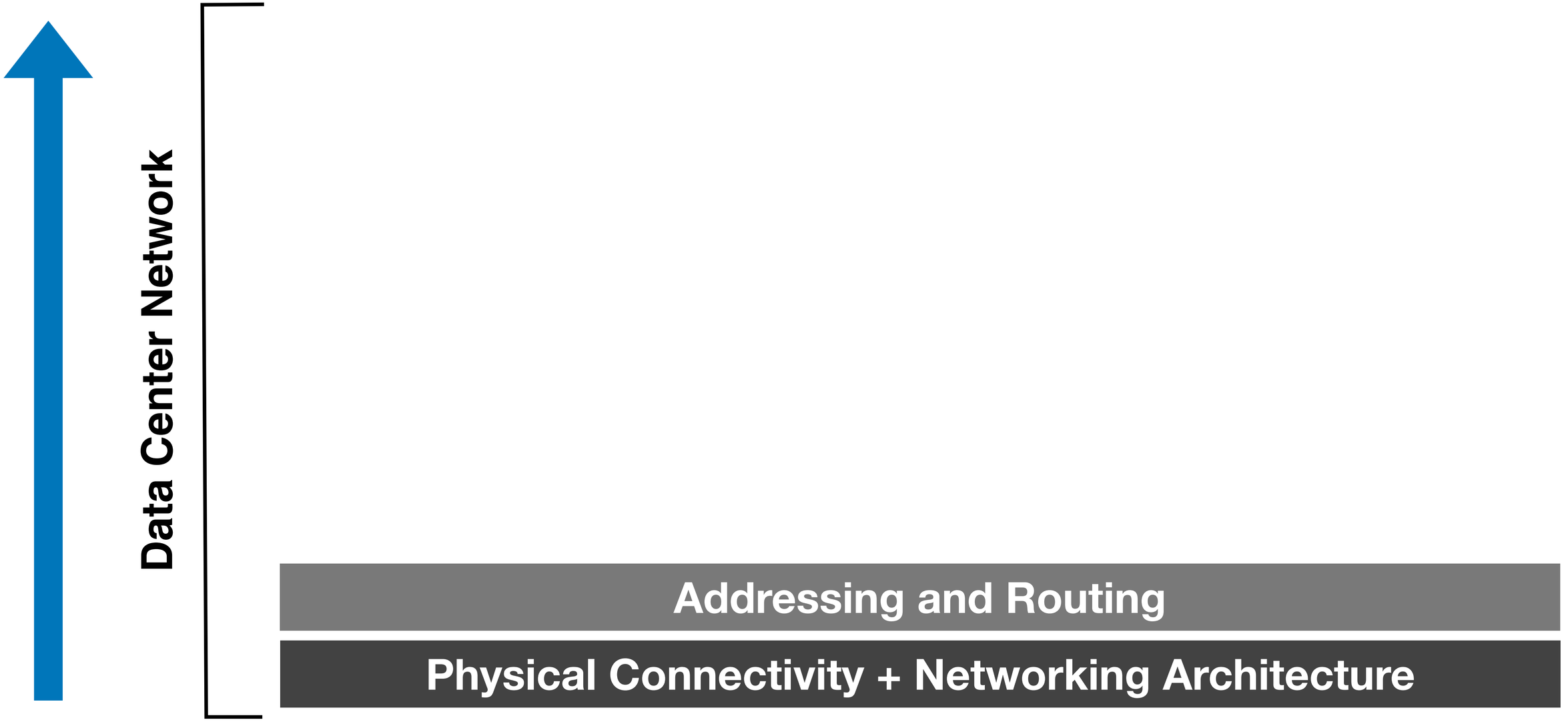


# Is this really different compared with what we have learned before?

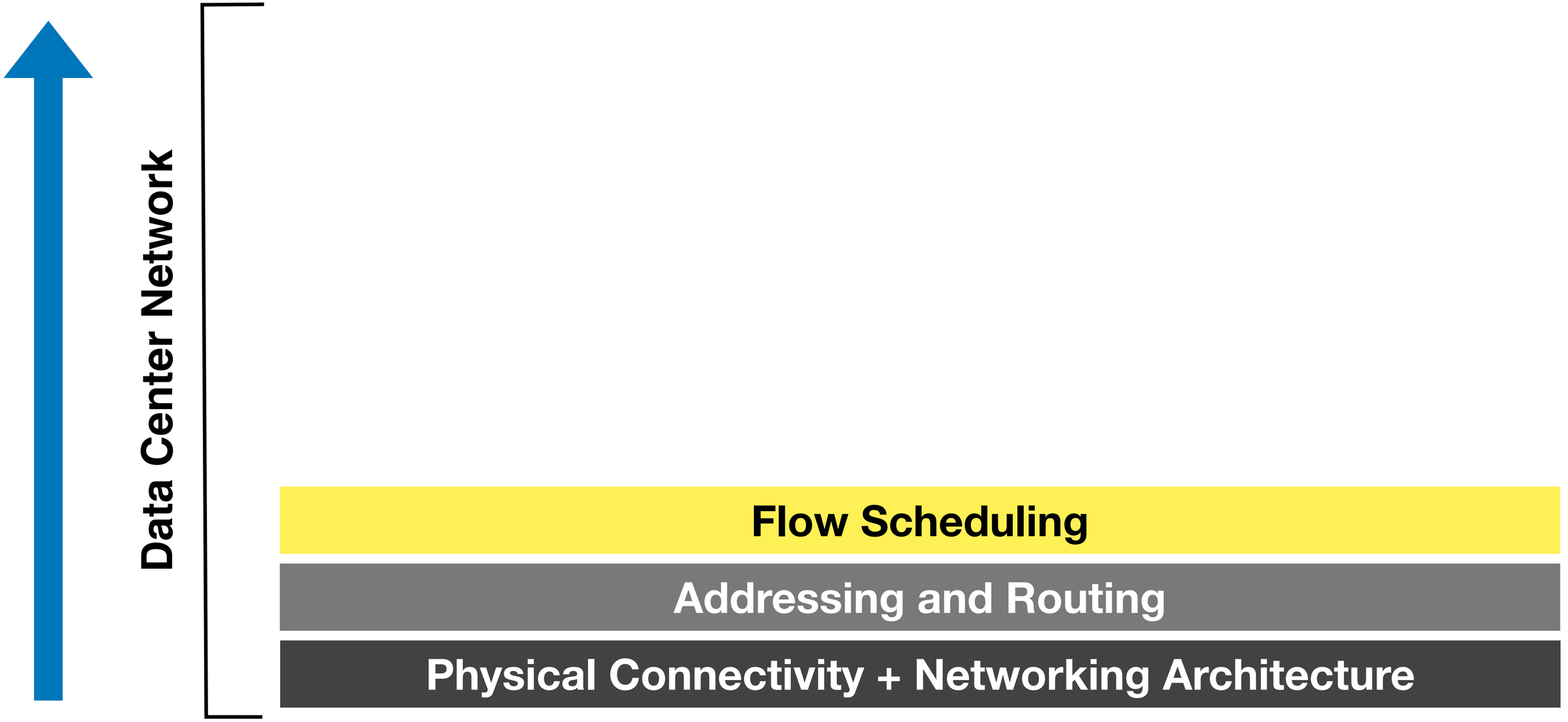
# Revisiting DCNet Layers for GPUs



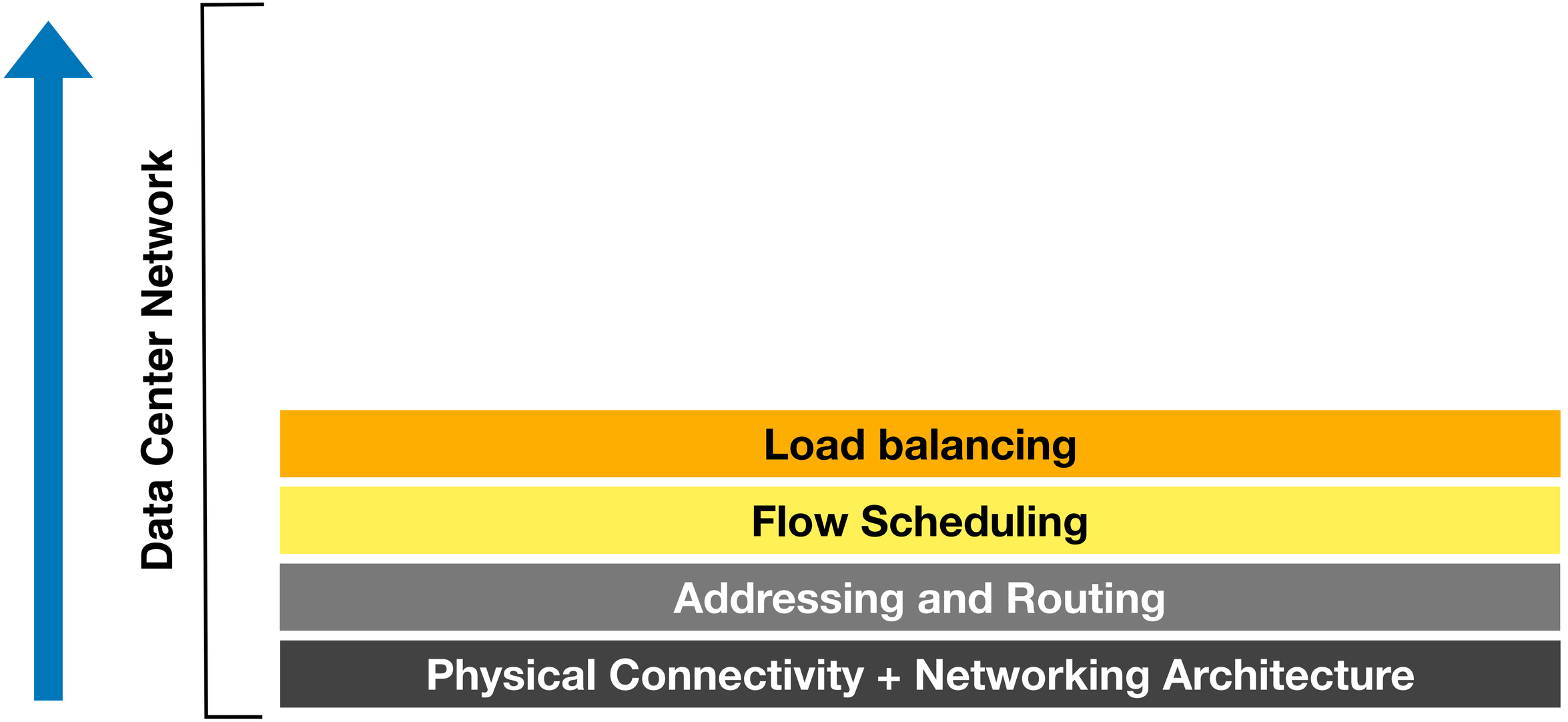
# Revisiting DCNet Layers for GPUs



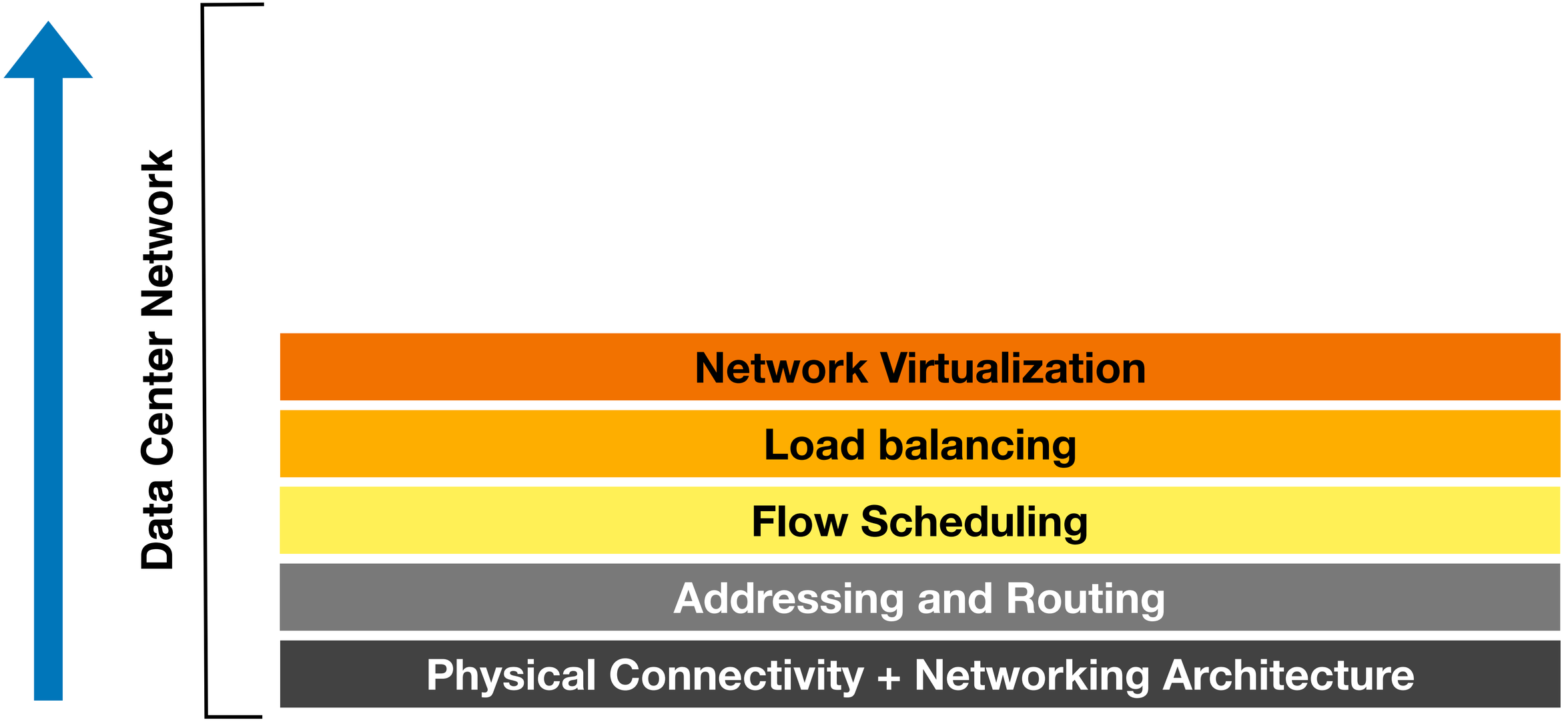
# Revisiting DCNet Layers for GPUs



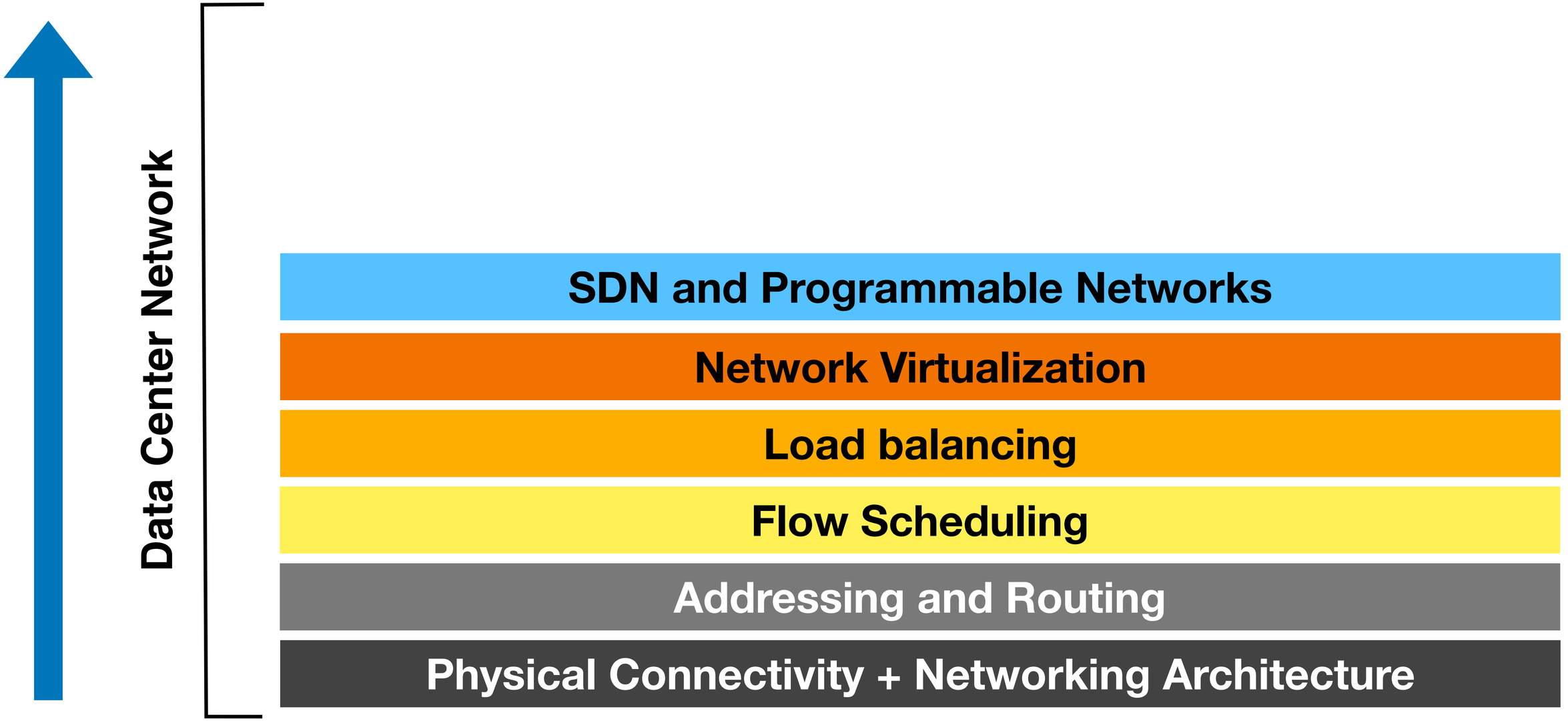
# Revisiting DCNet Layers for GPUs



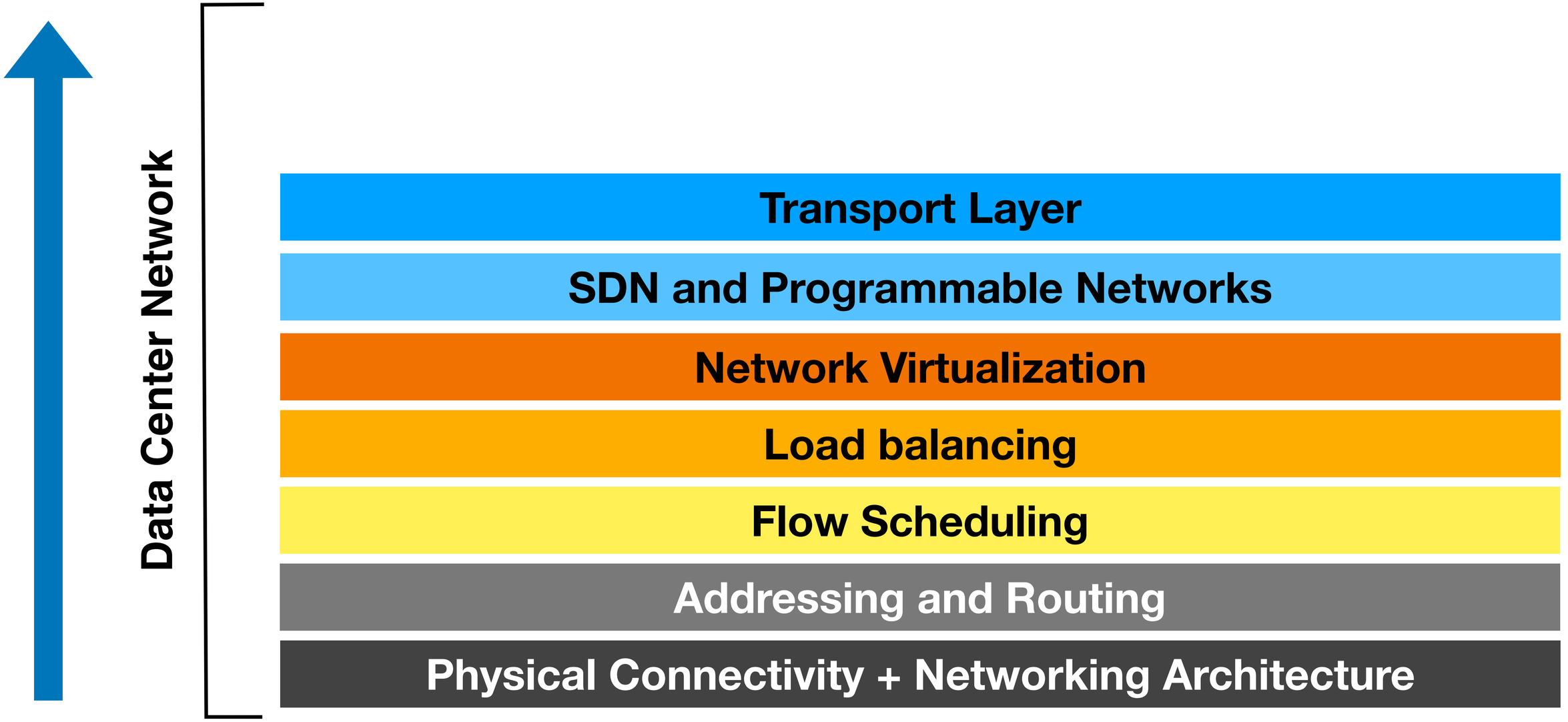
# Revisiting DCNet Layers for GPUs



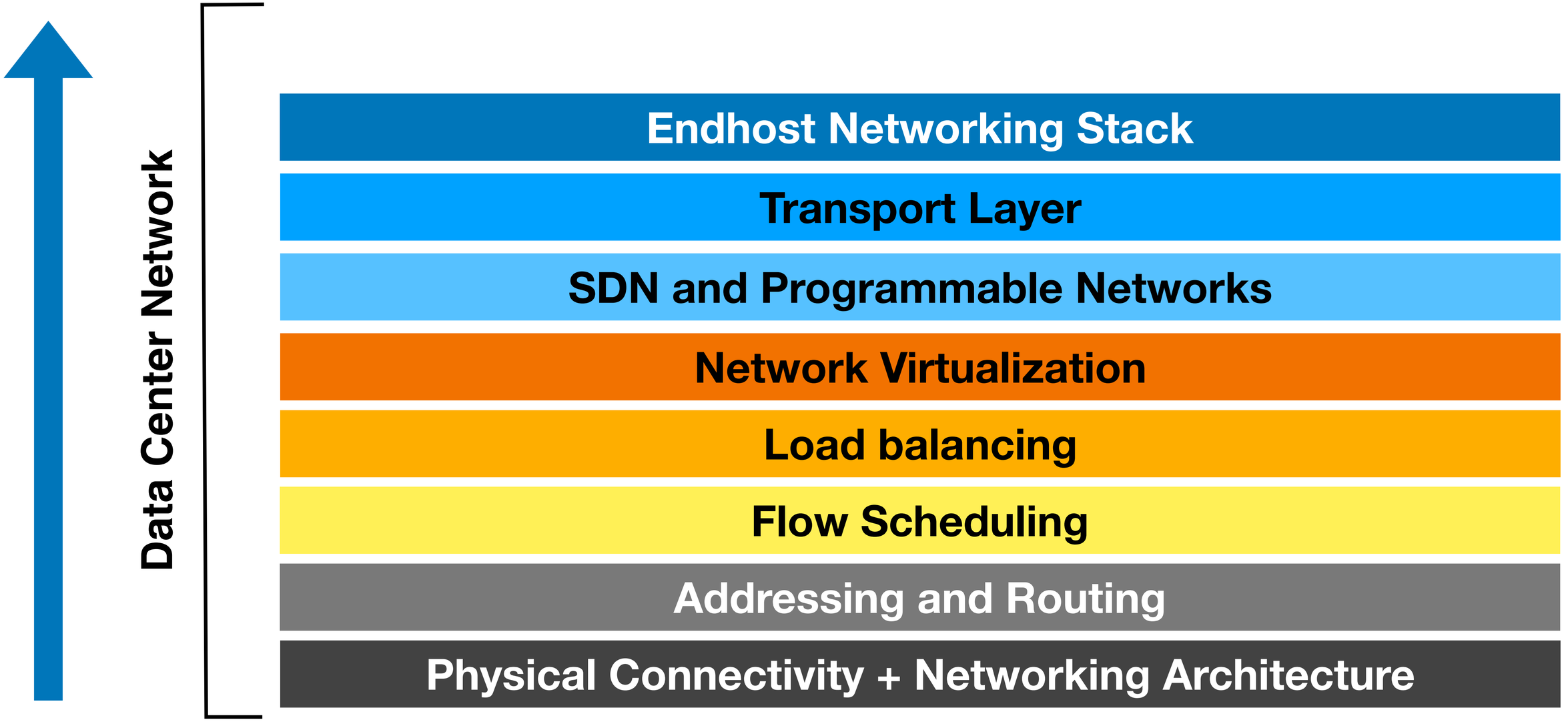
# Revisiting DCNet Layers for GPUs



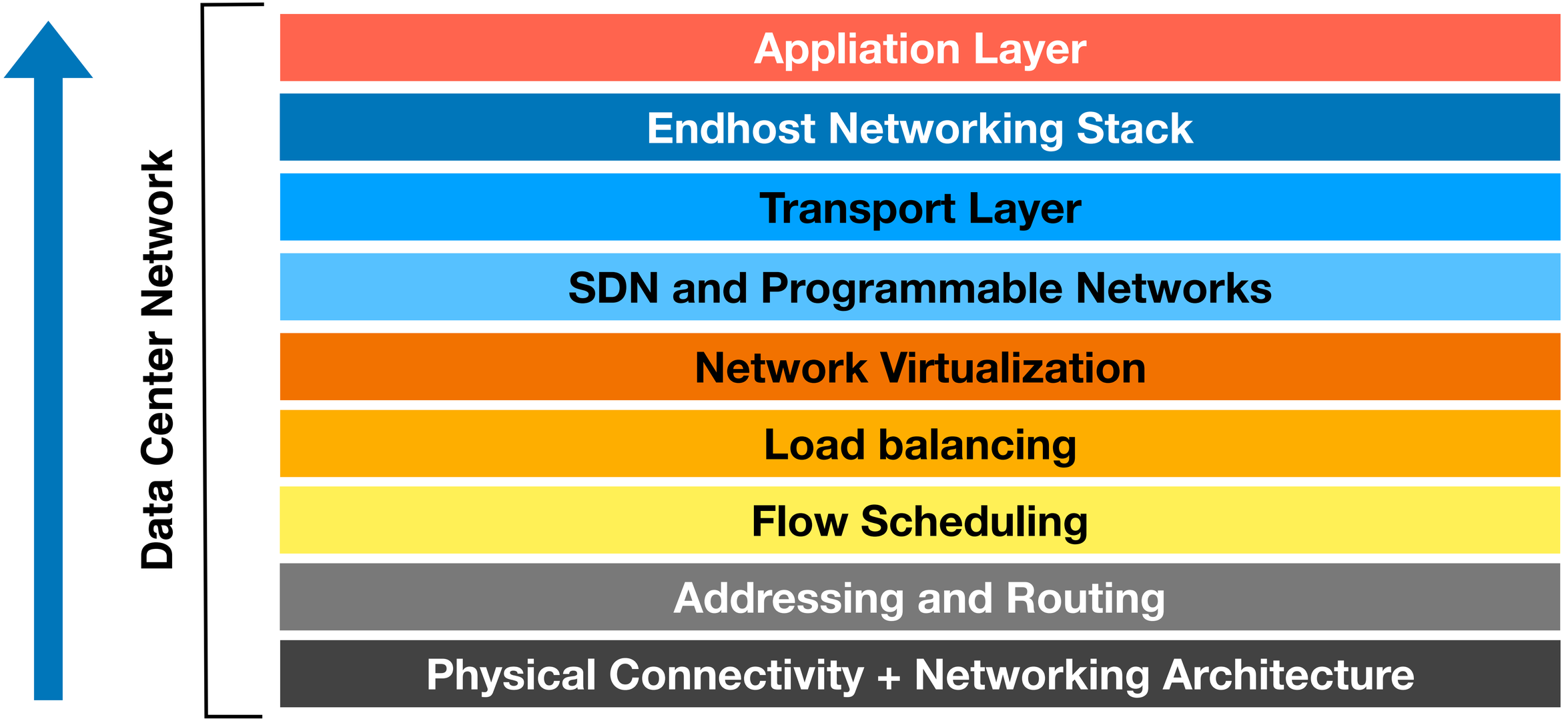
# Revisiting DCNet Layers for GPUs



# Revisiting DCNet Layers for GPUs



# Revisiting DCNet Layers for GPUs

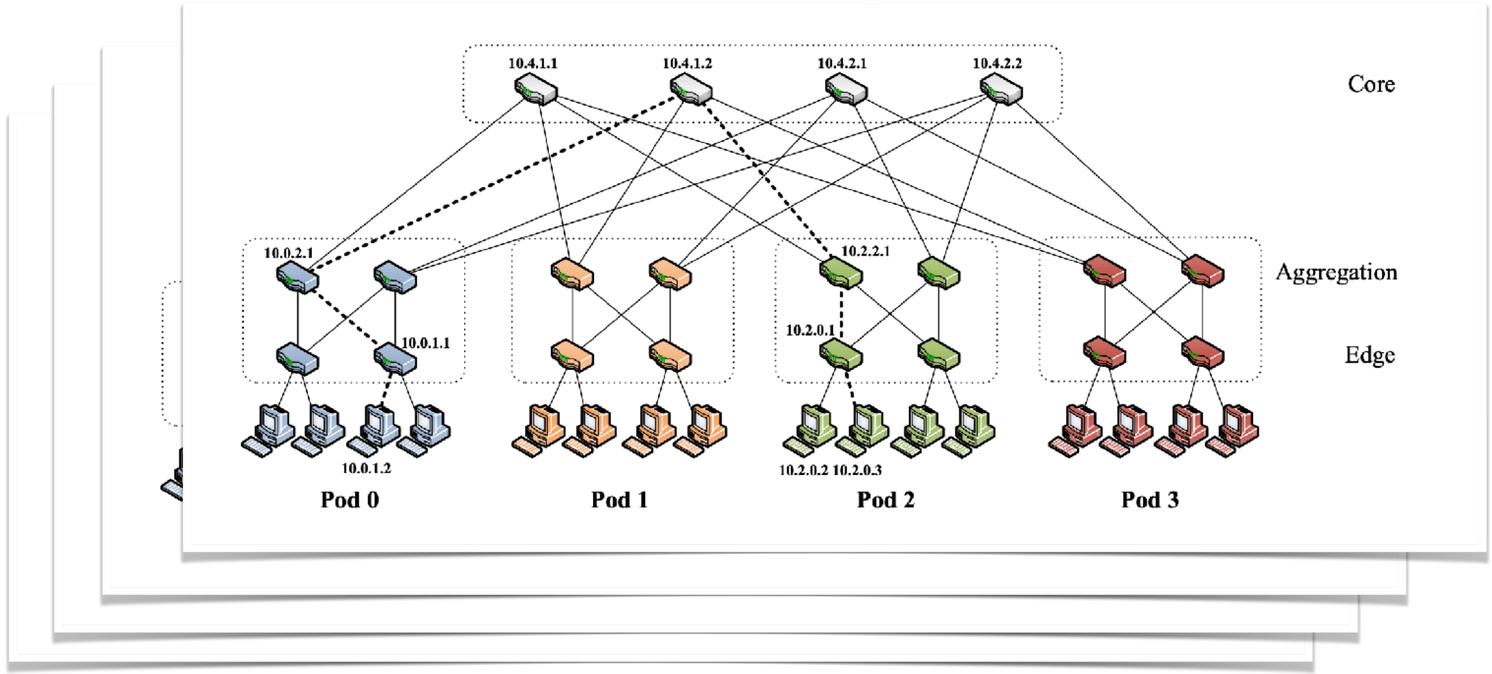


# Seems similar. So why don’t we reuse the existing DCNet?

# The Bandwidth Challenge!

# Multi-NICs —> Multiple Parallel DCNets

# Multi-NICs —> Multiple Parallel DCNets



# Summary

- Today
  - DCNet for GPUs (I)

- Next
  - DCNet for GPUs (II)