



# NVLink

AUTHOR: ARIAN AFZALZADEH

COURSE INSTRUCTOR: DR. FOSHATI

# Topics

1. MOTIVATION
2. NVLINK FEATURES
3. NVLINK ARCHITECTURE
4. NVLINK SIGNALING AND PROTOCOL
5. NVLINK GENERATIONS

# Motivation

The background is a light blue gradient. It features several 3D-rendered spheres in shades of blue and purple, some of which are partially cut off by the edges of the frame. At the bottom, there are larger, more complex organic shapes in similar colors, resembling liquid droplets or soft, inflated forms. The overall aesthetic is clean, modern, and minimalist.

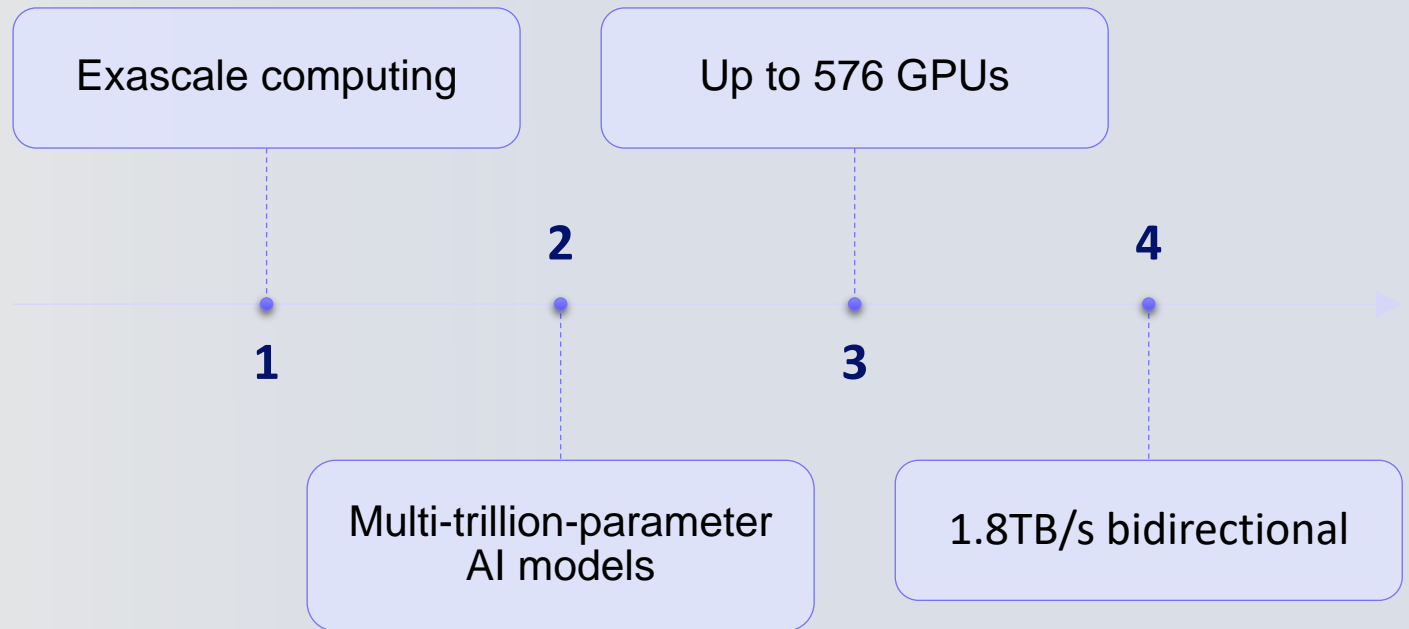
# Why NVLink Over PCIe?

- The need of a powerful interconnect in multiprocessing systems
- Achieve higher scalability for deep learning workloads
- Create an interconnect for GPUs that would offer much higher bandwidth than PCIe Gen 3
- Compatibility with the GPU ISA to support shared memory multiprocessing workloads (full support for Pascal's atomic operations)

# NVLink Achievements Today

Direct communication between every GPU within a server cluster

- High-bandwidth
- Energy-efficient
- Low-latency
- Lossless



The background features several abstract, 3D-rendered shapes in shades of blue and purple. On the left side, there are larger, more complex blobs and spheres. On the right side, there are smaller, simpler spheres. The overall aesthetic is clean and modern, with a focus on geometric forms and a cool color palette.

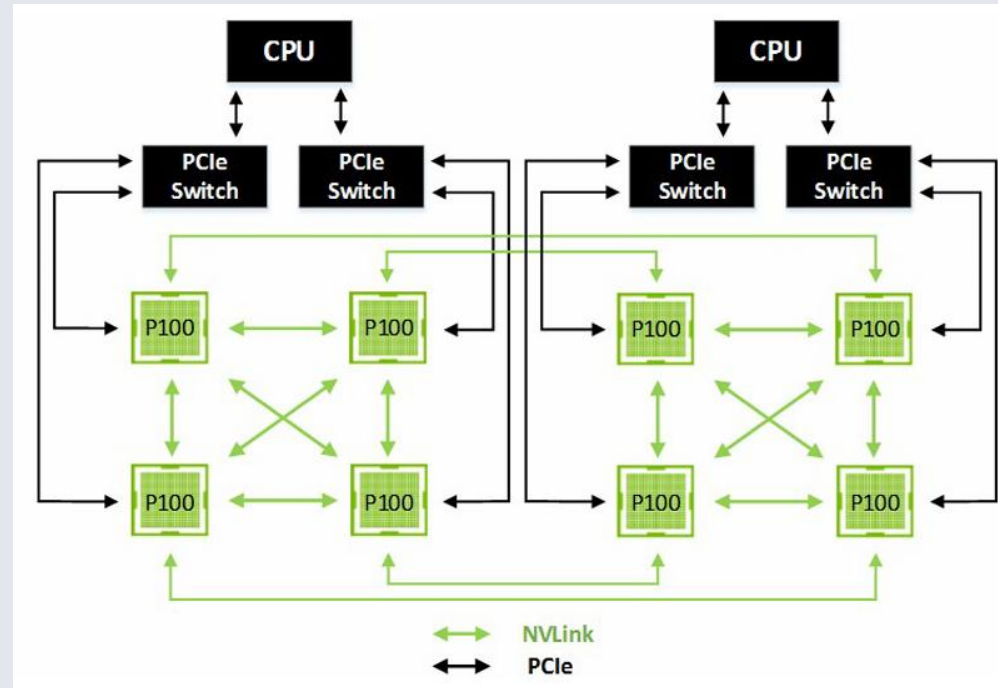
# NVLink Architecture

# First Generation of NVLink in The Tesla P100

- Bi-directional link at up to 40 Gb/s (GPU-to-GPU or GPU-to-CPU)
- Combination of four links for higher bandwidth

# GPU-to-GPU Connectivity

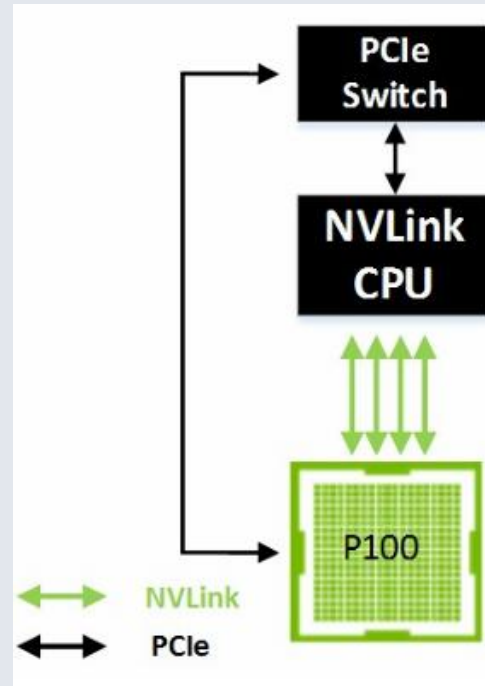
- Two fully NVLink-connected quads of GPUs, with NVLink connections between the quads
- GPUs within each quad connected to their respective CPUs through PCIe
- Less pressure on the PCIe uplink to CPUs and avoiding routing transfers through system memory



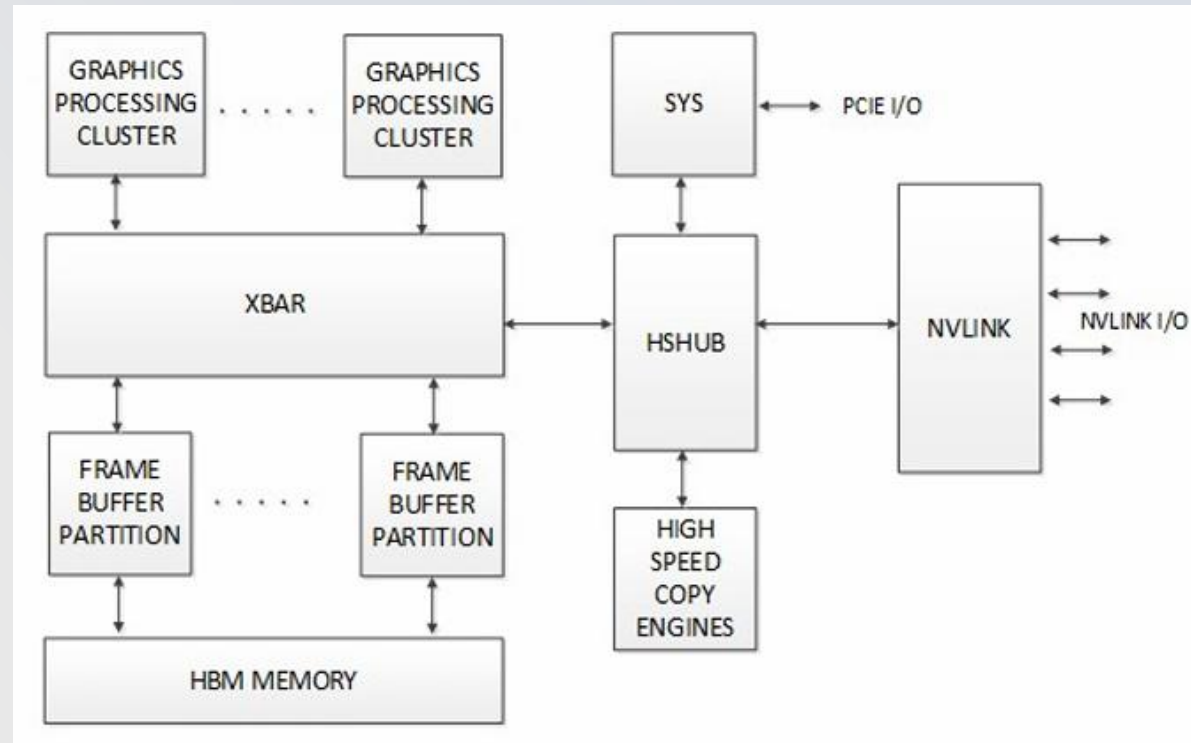


# CPU-to-GPU Connectivity

Access system memory at up to 160 GB/s bidirectional—5x higher than PCIe gen 3



# NVLink Relationship to Other Major Blocks in GP100



The background features a large, abstract, organic shape on the left side, rendered in shades of blue and purple with a soft gradient. This shape has several smaller, rounded protrusions. Scattered across the light blue background are numerous small, semi-transparent spheres in various sizes, also in shades of blue and purple, creating a bubbly, dynamic effect.

# **NVLink Signaling And Protocol**

# Signaling

- NVIDIA High-Speed Signaling interconnect (NVHS)
- Differential pair running at up to 20 Gb/s
- Combination of eight differential pairs in each direction to form a link
- Non-Return to-Zero
- Embedded clock
- Data sent from the PHY (physical level circuit) to the NVLink controller

# Physical Layer And Transaction Layer

The PL handles:

- Framing
- Scrambling/descrambling
- Polarity inversion
- Lane reversal
- Delivering the received data to the Data Link Layer

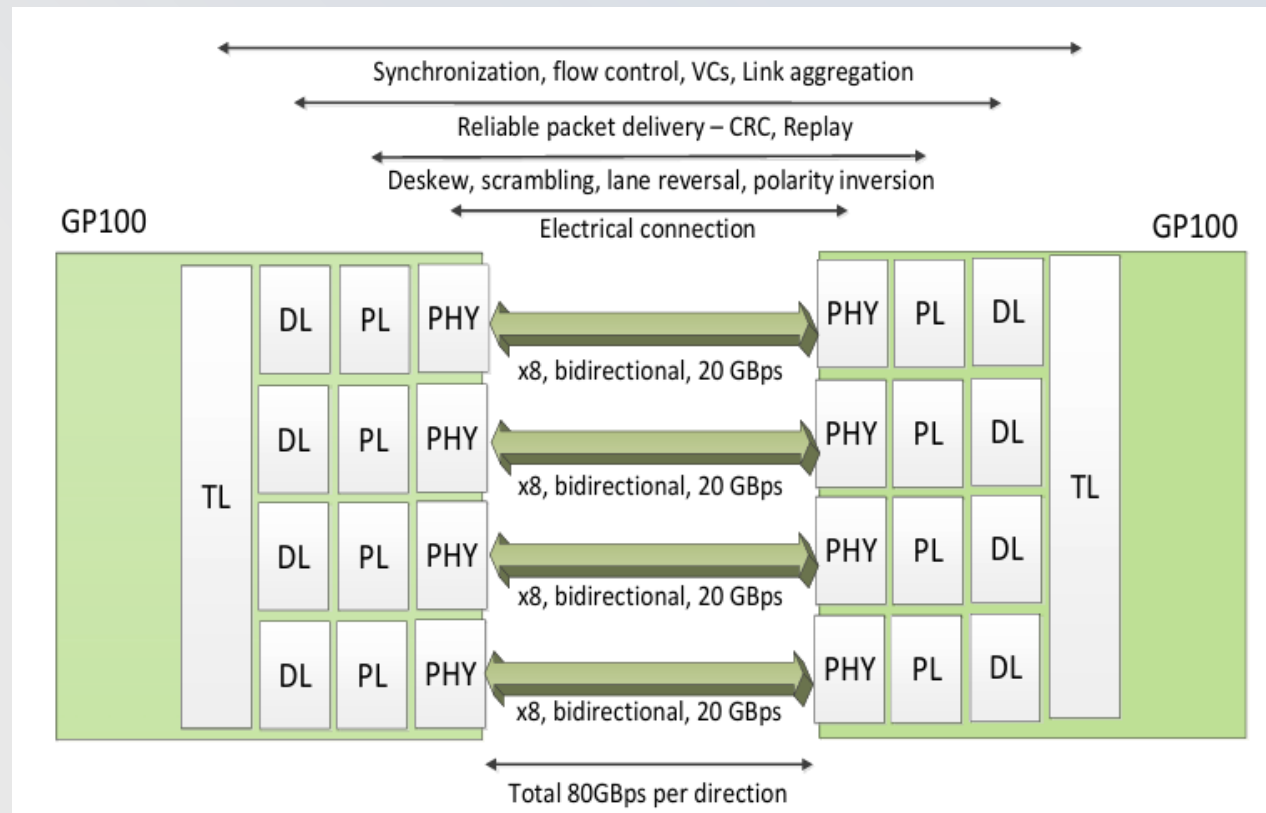
The TL handles:

- Synchronization
- Link flow control
- Aggregate multiple links together

# Data Link Layer

- Reliable transmission of packets across the link
- Protection of packets using a 25-bit CRC against up to 5 random bit errors or up to 25-bit bursts of errors
- Storing the transmitted packets in a replay buffer until they have been acknowledged by the receiver at the other end of the link
- On CRC error on incoming packet, retransmission of data from the replay buffer and no ACK
- A packet is retired from the replay buffer only when it has been acknowledged.
- Sending data on to the TL

# Overview of Layers And Links



# NVLink Generations

The background features a light blue gradient with various abstract elements. There are several spheres of different sizes in shades of blue and purple. A large, complex, organic shape in the lower center is composed of multiple rounded, overlapping forms in blue and purple. Other smaller spheres are scattered throughout the scene, some appearing as if they are floating or moving.



# Comparison of Different Generations

	Second Generation	Third Generation	Fourth Generation	Fifth Generation
NVLink bandwidth per GPU	300GB/s	600GB/s	900GB/s	1800GB/s
Maximum Number of Links per GPU	6	12	18	18
Differential Pairs per Link	8	4	2	2
Supported NVIDIA Architectures	Volta	Ampere	Hopper	Blackwell

The background features a light blue gradient with several 3D-rendered spheres in shades of blue and purple. On the right side, there are large, flowing, organic shapes in similar colors, creating a modern, abstract aesthetic.

# Thank you

ARIAN AFZALZADEH

ARIAN.AFZALZADEH01@SHARIF.EDU