DRIVENETS

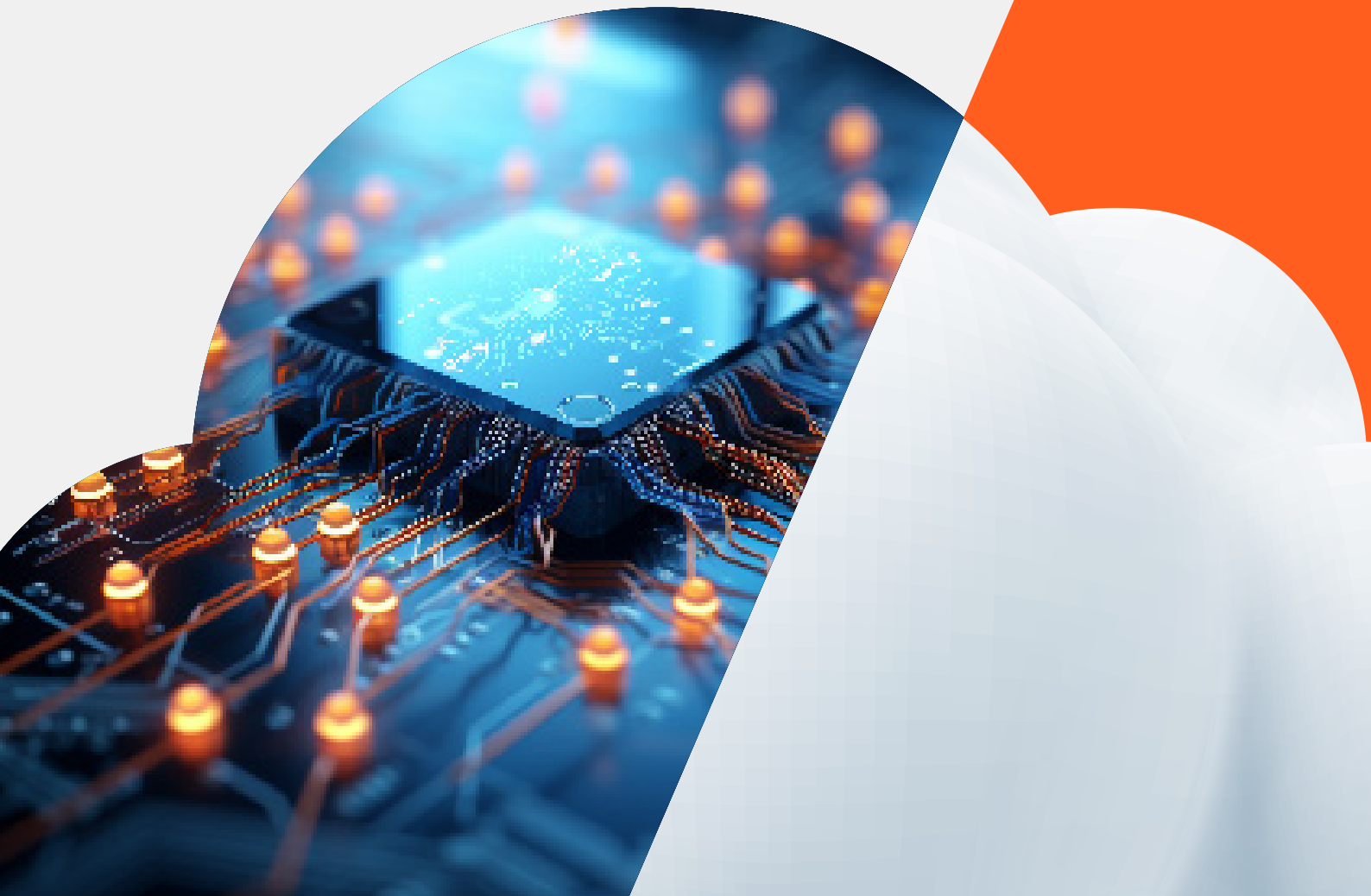# AI Cluster Reference Design

Guide for Building an 8K GPU Cluster with Network Cloud-AI

eGuide

Contents

Introduction

When building a large GPU cluster for artificial intelligence (AI) training purposes, the backend network fabric should be a high-performance, lossless and predictable one.

[DriveNets Network Cloud-AI](#) is a distributed network operating system (NOS). The solution manages a scheduled fabric supporting the highest performance AI networking at scale, with up to 32K GPUs (up to 800 Gbps each) per cluster, from any GPU vendor, with any Ethernet NIC.

This scheduled fabric AI backplane supports congestion-free operations, acting as a single Ethernet entity (with a single Ethernet hop from any GPU to any other GPU in the cluster) at any scale. It provides predictable, lossless connectivity with ultra-fast (microseconds level) failure recovery at scale. As a result, it delivers a proven >10% improvement in job completion time (JCT) for AI workloads at large-scale GPU clusters.

This document describes the capabilities of DriveNets Network Cloud-AI and showcase a high-level reference design for an 8,000 GPU cluster, equipped with 400Gbps Ethernet connectivity per GPU. This design explores network segmentation, high-performance fabrics, and scalable topologies, all optimized for the unique demands of large-scale AI deployments.

GPU cluster network architecture

An 8,000 high-performance GPU cluster utilizes a two-tier leaf-spine network topology designed for optimal performance and scalability in AI applications, leveraging the capabilities of DriveNets Network Cloud-AI.

The leaf tier consists of DriveNets NCP5-AI leaf switches powered by Broadcom Jericho3-AI processors, while the spine tier consists of DriveNets NCF2-AI spine switches powered by dual Broadcom Ramon3 processors. The NCP5-AI leaf switches ensure efficient communication between GPUs within a server rack, while the NCF2-AI spine switches, with their high-performance Broadcom Ramon3 processors, facilitate traffic routing across the entire cluster.
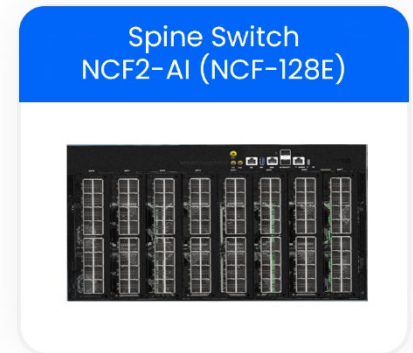
	Leaf Switch NCP5-AI (NCP-38E)	Spine Switch NCF2-AI (NCF-128E)
	Leaf: NCP5-AI	Spine: NCF2-AI

|		
Network Interfaces	18 x 800G OSFP	
Fabric Interfaces	20 x 800G / 40 x 400G OSFP	128 x 800G / 256 x 400G OSFP
Switching Capacity	30.4Tbps (14.4T Ethernet + 16T Cell Switching)	102.4T Cell Switching
Switching ASIC	Broadcom Jericho3-AI	2 x Broadcom Ramon3

Leaf tier: Each NCP5-AI switch directly connects to multiple GPU servers. These switches serve a capacity of 30.4Tbps. Each switch is equipped with 18 ports that can be configured for 800Gbps or 36 ports of 400Gbps speeds, offering flexibility based on specific GPU connectivity needs. NCP5-AI utilizes 20x800G cell-switching ports for high-bandwidth uplink connections to the spine layer. Notably, these 800G ports can be broken out into 40 interfaces of 400G.

Spine tier: These NCF2-AI switches handle traffic routing between leaf switches using cell-switching technology for efficient load balancing. NCF2-AI offers a massive capacity of 102.4Tbps with 128x800G cell-switching ports. Similar to the leaf tier, these ports can be broken out to 400G interfaces. This breakout functionality increases the spine radix, allowing connectivity for up to 256 leaf nodes. With full 800G utilization, this design supports up to 32K GPU connections within the cluster.

Example - Building an 8,192 GPU cluster

This reference design allows for a highly scalable network thanks to DriveNets Network Cloud-AI NOS and the breakout capabilities of NCP5-AI and NCF2-AI switches. The network can support up to 4,608 GPU connections with 800Gbps bandwidth in its base configuration. However, by leveraging 400G breakout on the leaf switches, one can connect up to 9,216 GPUs. (Typically, since only 32 [=8x4] of the 36 leaf ports are used, only 8,192 GPUs are connected.) Additionally, multitier spine functionality enables a network with up to 32K potential 800G GPU connections.

This section details how to construct a high-performance GPU cluster supporting 8,192 GPUs, each equipped with 400Gbps connectivity, leveraging the scalability of DriveNets Network Cloud-AI NOS and the breakout capabilities of its NCP5-AI leaf switches.
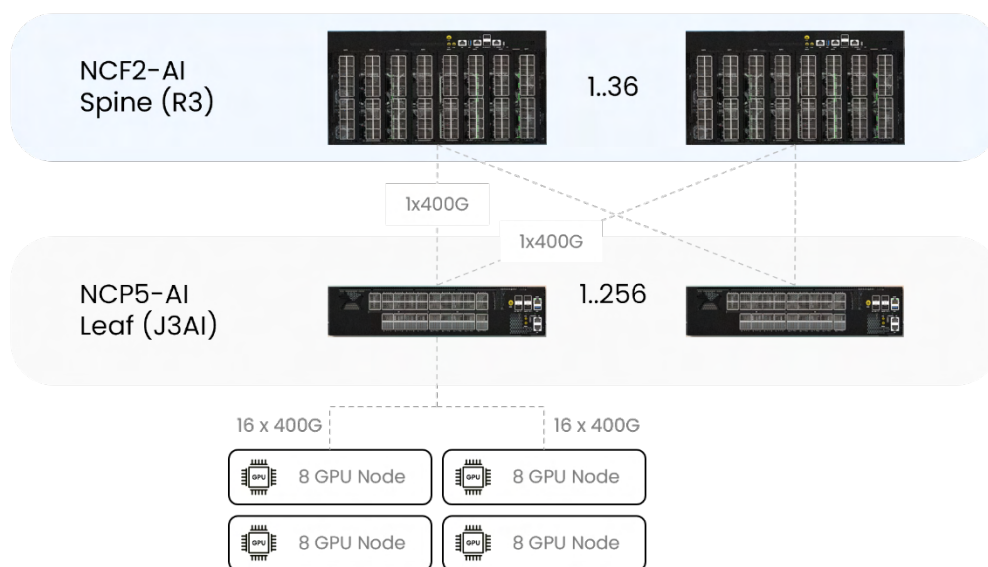
Network Cloud-AI: 8,192 x 400G GPU cluster

Network configuration

- **Leaf switches:** This design utilizes 256 DriveNets NCP5-AI leaf switches, each with a capacity of 30.4Tbps, specifically designed for AI workloads.

- **Leaf switch configuration:**
    - **GPU connectivity:** Each NCP5-AI switch directly connects to multiple GPUs. To maximize GPU attachment, we leverage 16 out of the available 18 ports with 800Gbps speeds on each leaf switch. These ports are broken out into 32x400G ports using DriveNets' breakout functionality. This configuration enables each leaf switch to support 32 GPUs with dedicated 400Gbps connections.

    - **Fabric connectivity:** While 32 uplink ports per leaf switch would suffice for GPU traffic, we utilize 36 available fabric ports for redundancy (N+1) and to accommodate additional traffic from Broadcom's internal cell control packets. This translates to 32 ports dedicated to GPU traffic and 4 additional ports for redundancy and cell overhead.

- **Spine switches:** The number of required spine switches is determined by the total uplink capacity from the leaf switches.

Spine switch calculations

- **Total uplink capacity:** We calculate the total uplink capacity by multiplying the number of leaf switches (256) by the number of fabric ports per leaf switch (36): 256 leaves * 36 ports/leaf = 9,216 uplink ports (400Gbps).

- **Number of spines:** To handle this uplink traffic, we require 36 spine switches. This calculation is derived by dividing the total uplink ports (9,216) by the number of connections a single spine can support from a single leaf (256 x 400Gbps breakout ports per spine).

Rack elevation and data center layout

While data center deployments for GPU clusters can vary in design, key factors like power consumption and rack size play a crucial role. This reference design adopts a conservative approach, assuming a maximum power draw of 25kW per rack and utilizing a standard 48-unit rack size. This configuration aligns with common practices for high-performance GPU hosts like the NVIDIA HGX H100.
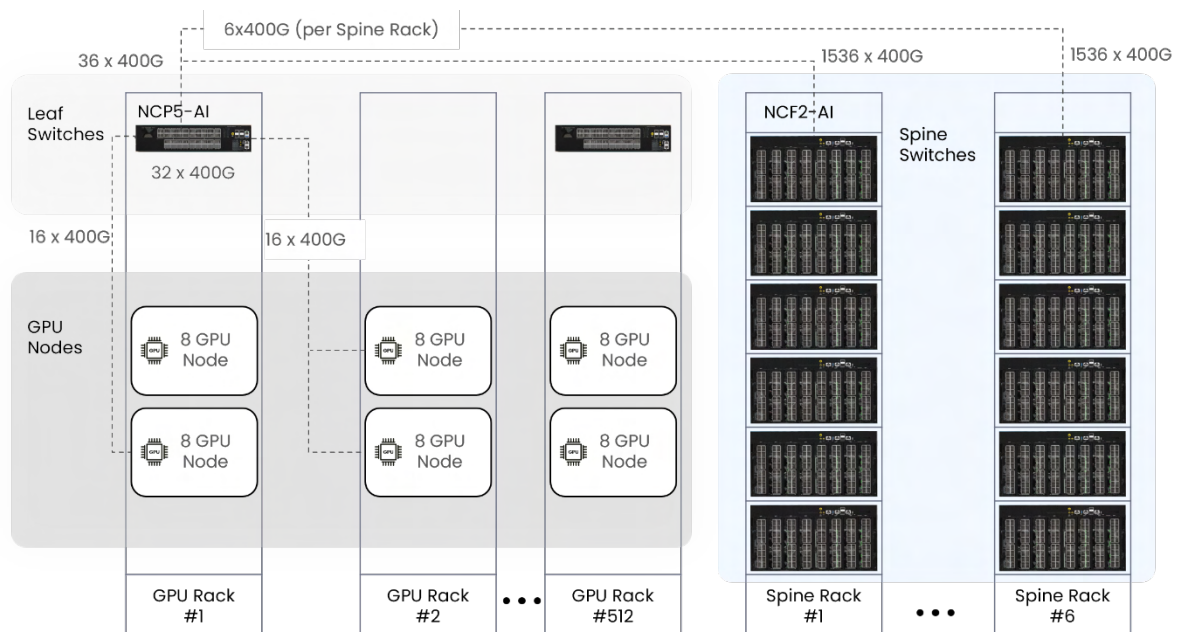
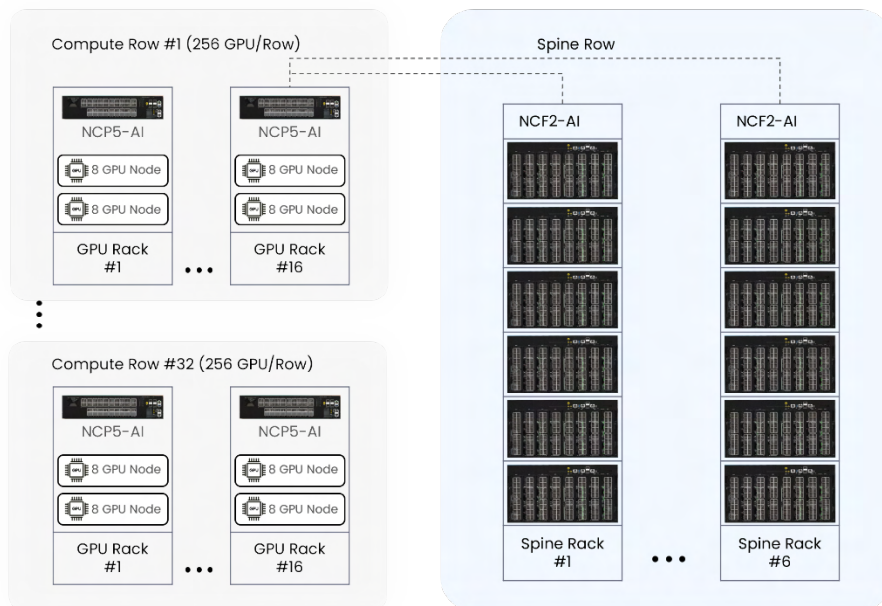*Network Cloud-AI: 8,192 x 400G GPU rack elevation*

Network components

The solution comprises four key components:

1. **Compute racks:** These racks house the computational workhorses – the GPU nodes and the leaf switches. In this reference design, we consider a configuration with 2 GPU nodes per rack, each node supporting 16 Ethernet-capable GPUs (8 GPUs per node). This assumption aligns with typical data center practices, where a high-performance host like the NVIDIA HGX H100 might require around 11kW of power. Considering network and management switches, the entire rack's power consumption would range from 22kW to 25kW. Since each

leaf will support 32x400G GPUs, every two compute racks will connect to a single DriveNets NCP5-AI leaf switch.

2. **Spine racks:** These racks house the spine switches. Each spine rack can contain 6 DriveNets NCF2-AI spine switches. These racks are usually located in a central location in the data center as all compute racks which house the GPU nodes and the leaf switches need to have connectivity to each of the spines in the network.

3. **Host NIC-to-leaf connectivity:** Connectivity of GPU to leaf switch within the rack and between racks can utilize cost-effective direct attach copper (DAC) cables instead of optical fibers, leading to power and cost savings. In this design, a single NCP5-AI leaf will have 16x400G in-rack connections and 16x400G connections to the neighbor rack.

4. **Leaf-to-spine connectivity:** Each rack equipped with a leaf switch will have 36 of its 400G links dedicated for connections to the spine switches. These spine switches reside in separate spine racks, typically designed to accommodate 48 rack units (RU). Each spine rack can comfortably house 6 DriveNets NCF2-AI spine switches. This design ensures that each leaf switch connects to every spine switch using a single 400Gbps link, resulting in a total of 6x400G connections per leaf switch to each spine rack.

Network Cloud-AI: 8,192 x 400G GPU data center layout

Reference Architecture

800G GPU ports	400G GPU ports	Num of Leaf (NCP)	Num of Spine (NCF)	Fabric Architecture
400G Platforms				
	128	4	3	Two-Stage
	512	16	12	Two-Stage
	1,536	48	36	Two-Stage
800G Platforms				
1,024	2,048	64	8	Two-Stage
2,048	4,096	128	18	Two-Stage
4,096	8,192	256	36	Two-Stage
8,192	16,384	512	216 (144+72)	Three-Stage
16,384	32,000	1,024	432 (288+144)	Three-Stage
32,000	–	2,048	864 (576+288)	Three-Stage

- 32,000 is the maximum available number of Virtual Output Queue (VOQ) per AI Fabric

- Network to Fabric speedup ratio is 10% on NCP leaf. (e.g 16 Network ports and 18 Fabric ports)

400G Platforms

NCF1 (NCF-48CD)
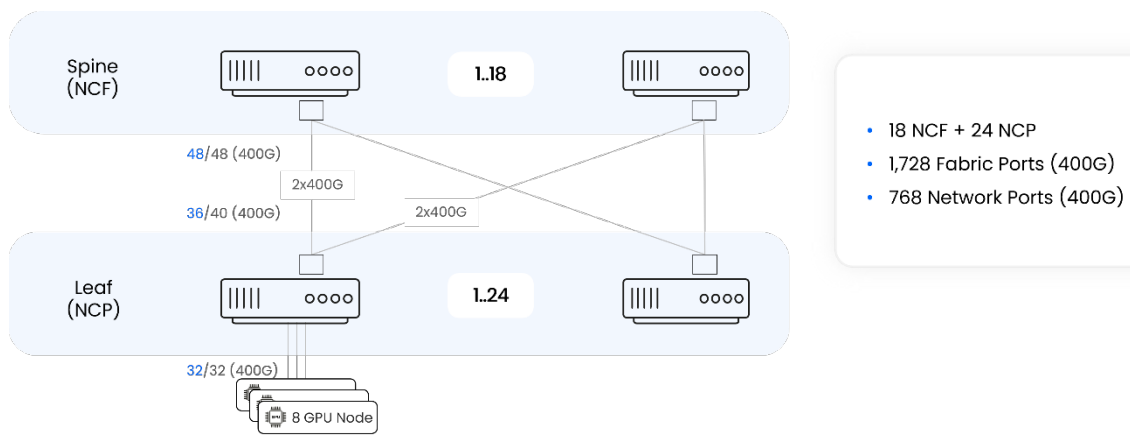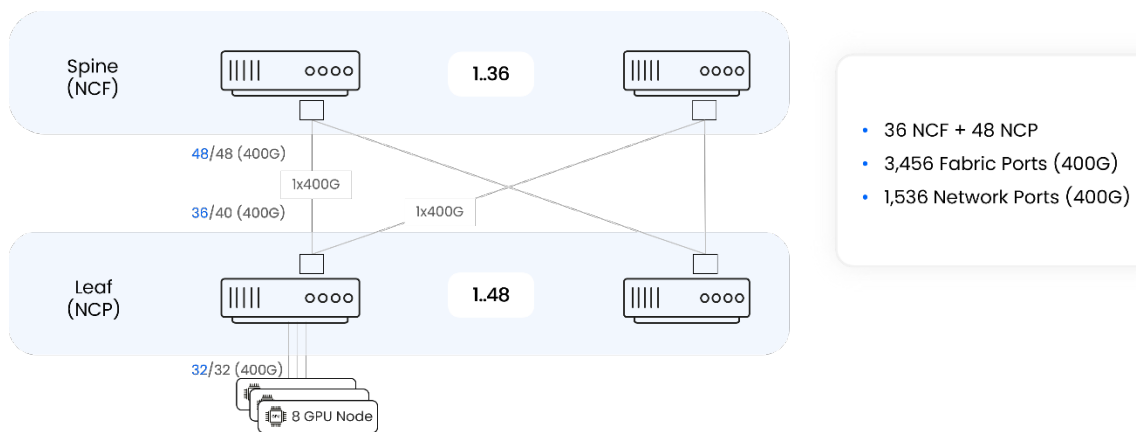

2x R1

48x4000G (19.2T)

NCP3 (NCP-36CD)

2xJ2C+

72x400G (28.8T)

32x400GE + 40x400GE Fabric

768 x 400G AI Fabric

Spine (NCF)   1..18

48/48 (400G)
2x400G
36/40 (400G)   2x400G

Leaf (NCP)   1..24

32/32 (400G)

8 GPU Node

- 18 NCF + 24 NCP
- 1,728 Fabric Ports (400G)
- 768 Network Ports (400G)

1,500 x 400G AI Fabric

Spine (NCF)   1..36

48/48 (400G)
1x400G
36/40 (400G)   1x400G

Leaf (NCP)   1..48

32/32 (400G)

8 GPU Node

- 36 NCF + 48 NCP
- 3,456 Fabric Ports (400G)
- 1,536 Network Ports (400G)
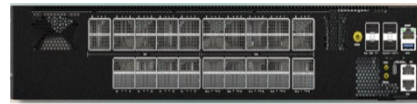
800G Platforms

NCF2 (NCF-128E)

2x R3

128x800G (102.4T)

NCP5-AI (NCP-38E)

J3AI

38x800G (30.4T)
18x800GE + 20x800GE Fabric

---

128 x 800G / 256 x 400G AI Fabric

Spine (NCF) — 1..2

72/128 (800G)
9x800G
18/20 (800G)
9x800G

Leaf (NCP) — 1..8

16/18 (800G)

8 GPU Node

- 2 NCF + 8 NCP
- 288 Fabric Ports (800G)
- 256 Network Ports (400G)

1K 800G / 2K 400G AI Fabric

Spine (NCF) — 1..9

128/128 (800G)
2x800G
18/20 (800G)
2x800G

Leaf (NCP) — 1..64

16/18 (800G)

8 GPU Node

- 9 NCF + 64 NCP
- 2,304 Fabric Ports (800G)
- 1,024 Network Ports (800G)

2K 800G / 4K 400G AI Fabric

Spine (NCF)

1..18

128/128 (800G)

1x800G

18/20 (800G)

1x800G

Leaf (NCP)

1..128

16/18 (800G)
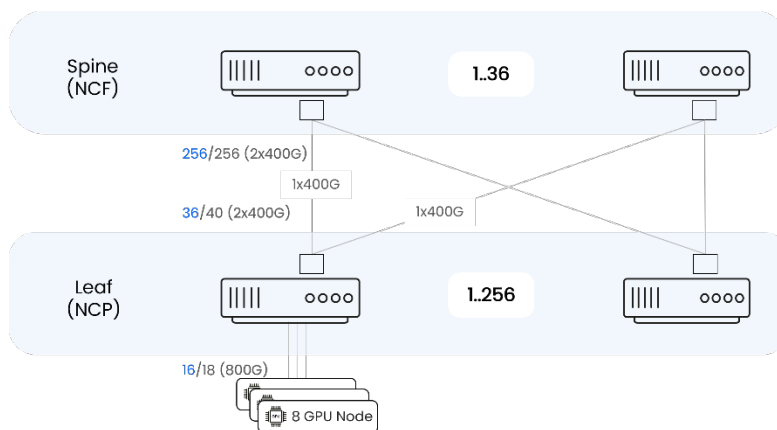
8 GPU Node
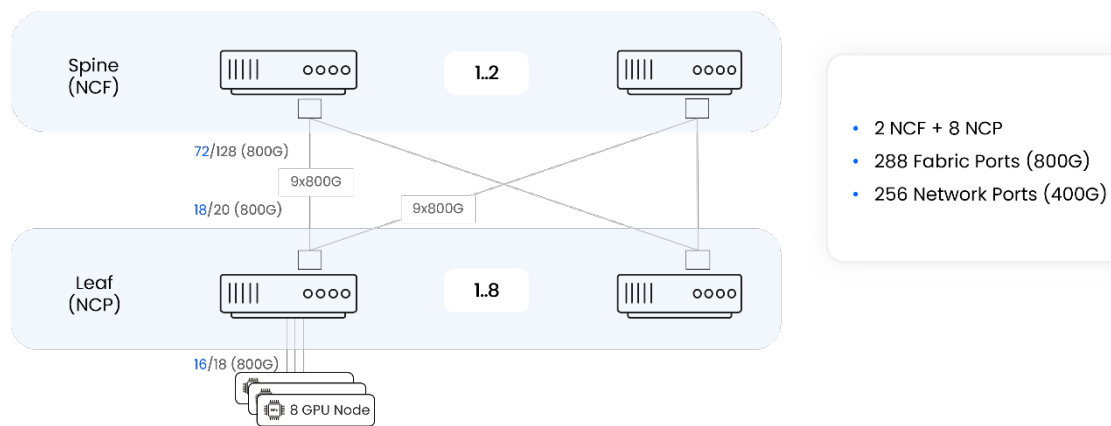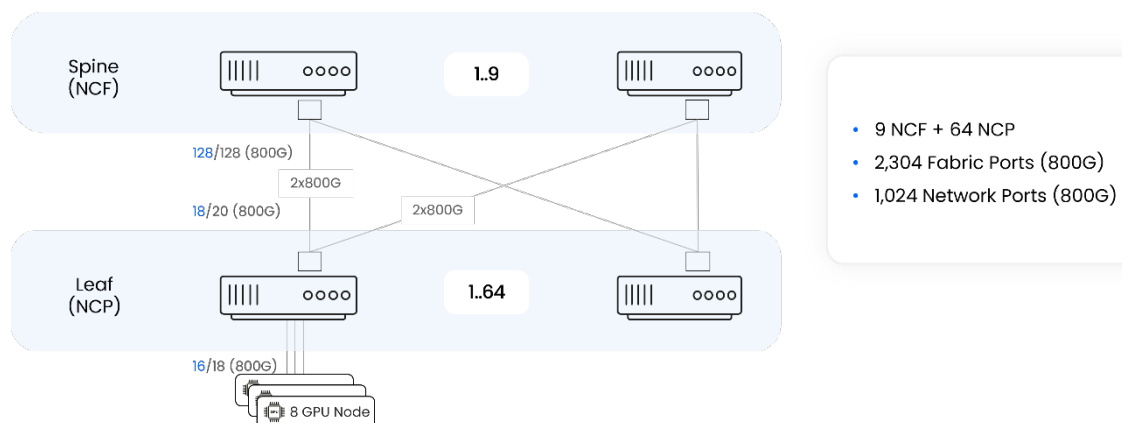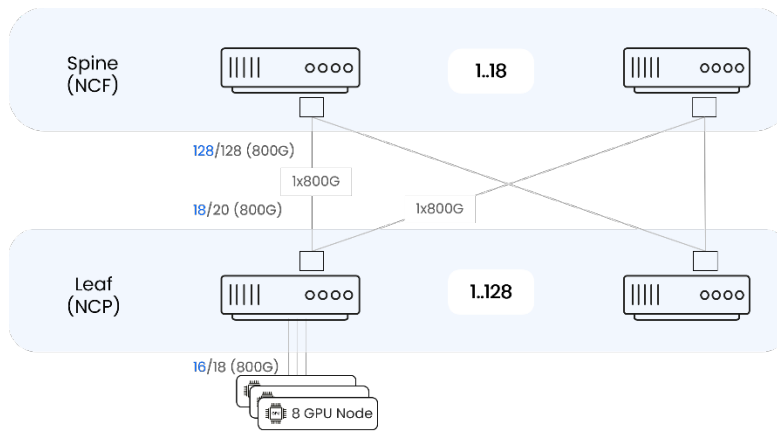
- 18 NCF + 128 NCP
- 4,608 Fabric Ports (800G)
- 2,048 Network Ports (800G)

4K 800G / 8K 400G AI Fabric

Spine (NCF)

1..36

256/256 (2x400G)

1x400G

36/40 (2x400G)

1x400G

Leaf (NCP)

1..256

16/18 (800G)

8 GPU Node

- 36 Spine + 256 Leaf
- 9,216 Fabric Ports (800G)
- 4,096 Network Ports (800G)

8K 800G / 16K 400G AI Fabric

Virtual Spine Domain
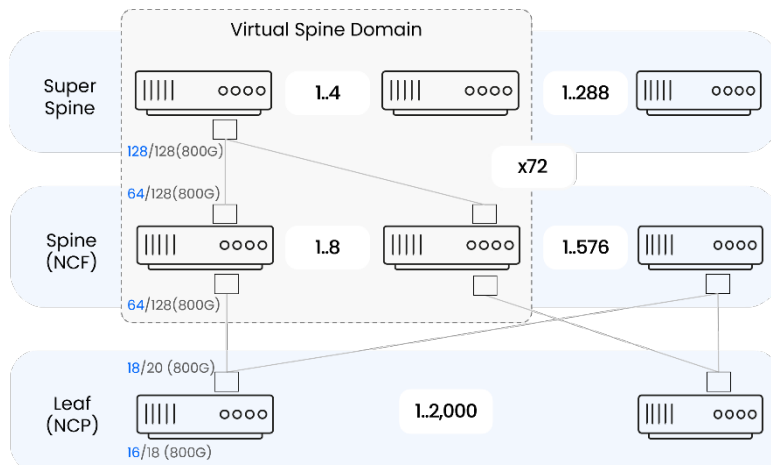
Super Spine

1..4

1..72

128/128(800G)

x18

64/128(800G)

Spine (NCF)

1..8

1..144

64/128(800G)

18/20 (800G)

Leaf (NCP)

1..512

16/18 (800G)
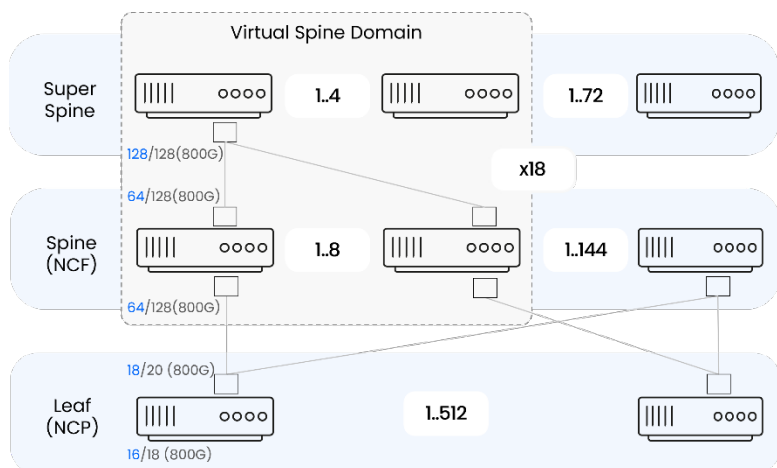
- 216 Spine + 512 Leaf
- 36,864 Fabric Ports (800G)
- 8,192 Network Ports (800G)

16K 800G / 32K 400G AI Fabric

Virtual Spine Domain

Super Spine   1..4   1..144

128/128(800G)

64/128(800G)

Spine (NCF)   1..8   1..288

x32

64/128(800G)

18/20 (800G)

Leaf (NCP)   1..1,024

16/18 (800G)
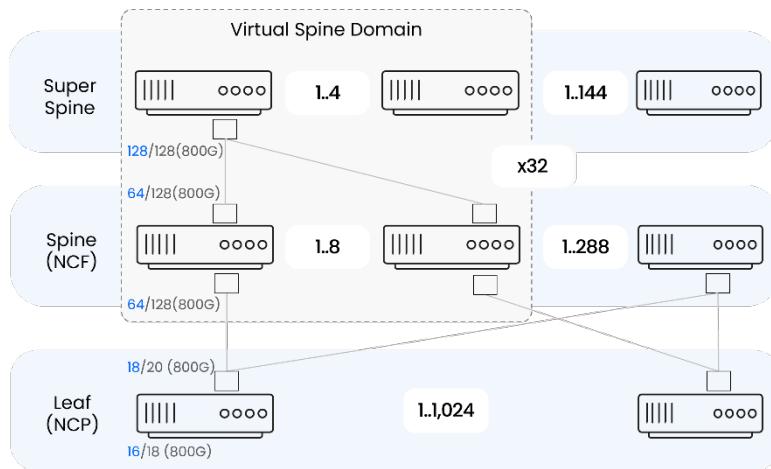
- 432 Spine + 1,024 Leaf
- 73,728 Fabric Ports (800G)
- 16,384 Network Ports (800G)

32K 800G AI Fabric

Virtual Spine Domain

Super Spine   1..4   1..288

128/128(800G)

64/128(800G)

Spine (NCF)   1..8   1..576
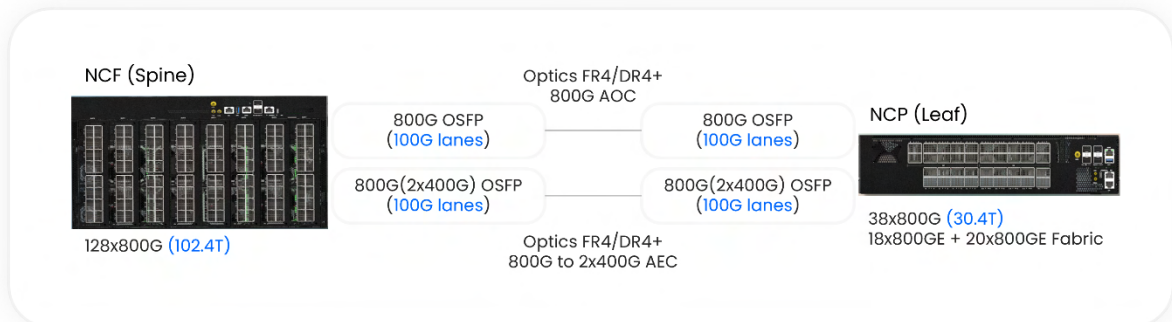
x72

64/128(800G)
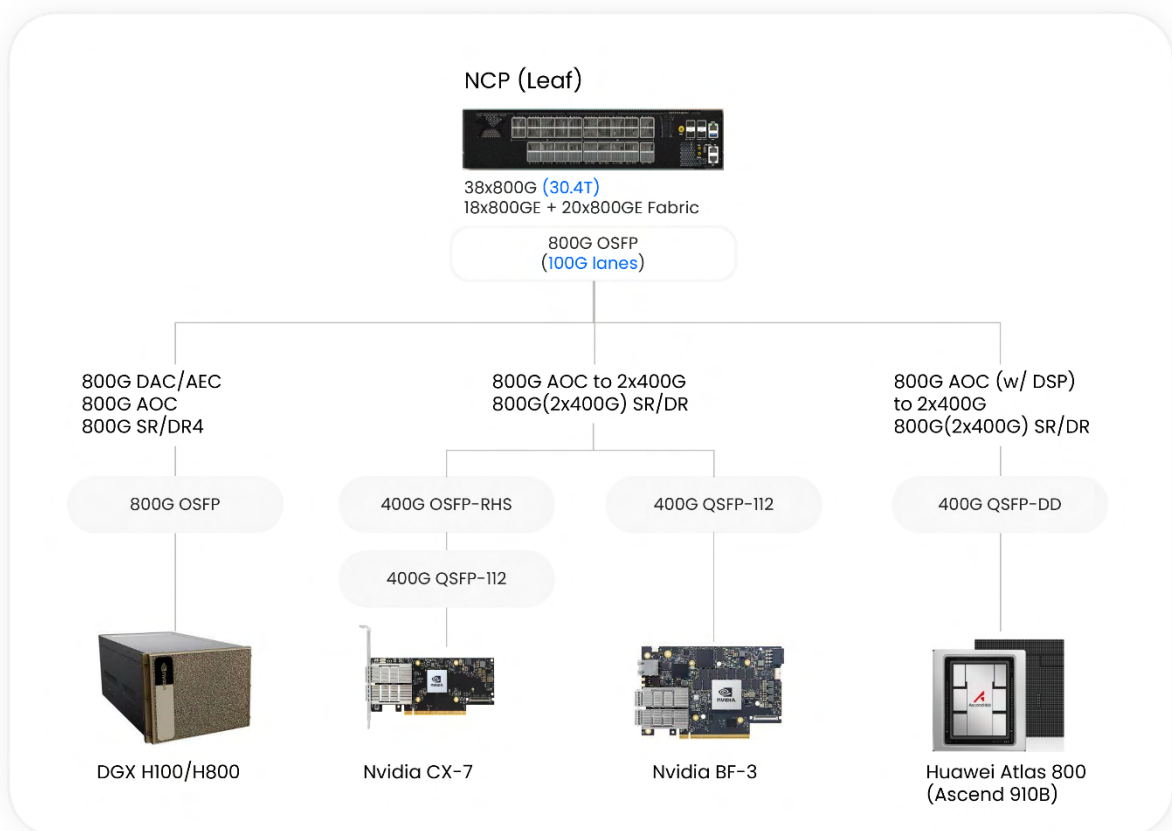
18/20 (800G)

Leaf (NCP)   1..2,000

16/18 (800G)

- 864 Spine + 2,000 Leaf
- 146,592 Fabric Ports (800G)
- 32,000 Network Ports (800G)

Cables and Optics

Leaf to Spine

NCF (Spine)

Optics FR4/DR4+
800G AOC

800G OSFP
(100G lanes)

800G OSFP
(100G lanes)

NCP (Leaf)

800G(2x400G) OSFP
(100G lanes)

800G(2x400G) OSFP
(100G lanes)

128x800G (102.4T)

38x800G (30.4T)
18x800GE + 20x800GE Fabric

Optics FR4/DR4+
800G to 2x400G AEC

Leaf to node

NCP (Leaf)

38x800G (30.4T)
18x800GE + 20x800GE Fabric

800G OSFP
(100G lanes)

800G DAC/AEC
800G AOC
800G SR/DR4

800G AOC to 2x400G
800G(2x400G) SR/DR

800G AOC (w/ DSP)
to 2x400G
800G(2x400G) SR/DR

800G OSFP

400G OSFP-RHS

400G QSFP-112

400G QSFP-DD

400G QSFP-112

DGX H100/H800

Nvidia CX-7

Nvidia BF-3

Huawei Atlas 800
(Ascend 910B)

Summary

This reference design leverages DriveNets Network Cloud-AI and the breakout capabilities of NCP5-AI leaf switches to create a highly scalable network foundation for an 8,192 GPU cluster. Each GPU benefits from a dedicated 400Gbps connection for efficient communication. The 256 leaf switches connect to 36 spine switches, ensuring redundancy and efficient traffic routing across the entire cluster.

This architecture can offer multiple ways to achieve low total cost of ownership (TCO). One way is by using copper instead of optics between leaf switches and GPU servers. The second and more important way is by reducing dramatically the end-to-end job completion time.

DRIV⚡NETS