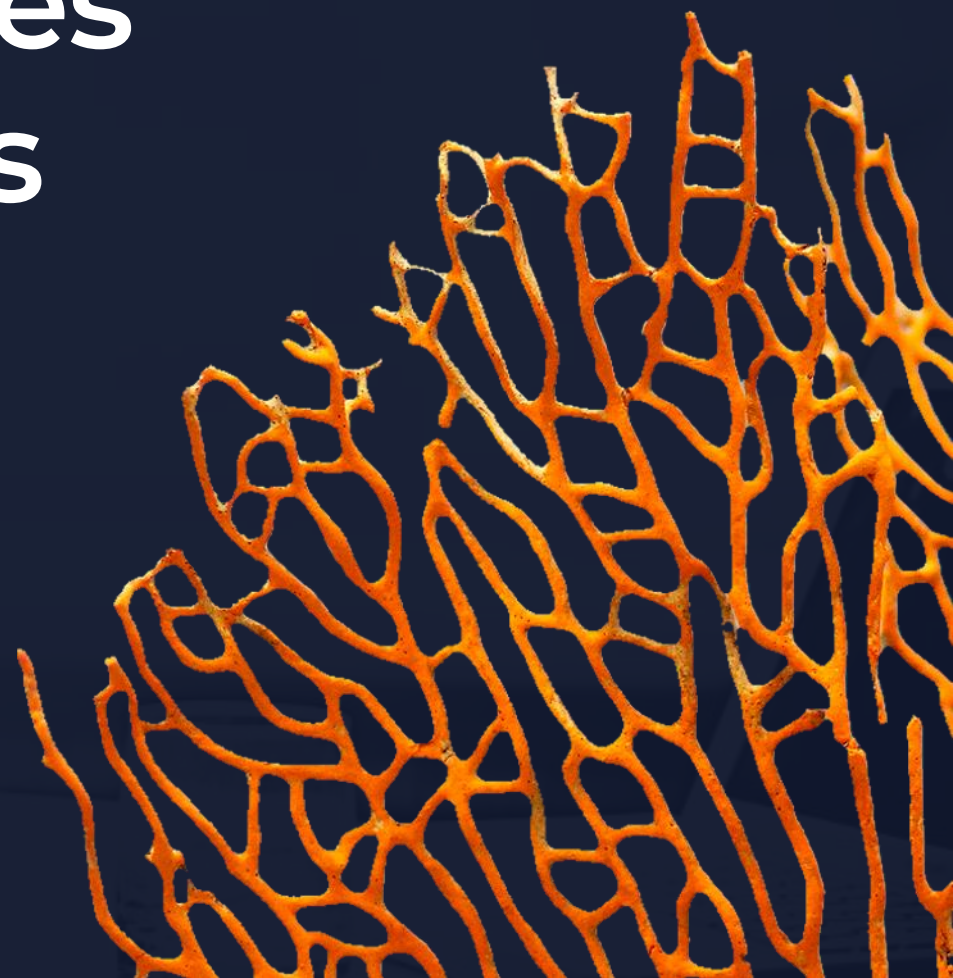


——> Murilo Boratto

Evolução dos Ambientes Supercomputacionais Brasileiros

29 DE JANEIRO
8H

Escola Supercomputador
Santos Dumont 2026
SD03-II



Supercomputador
Santos Dumont

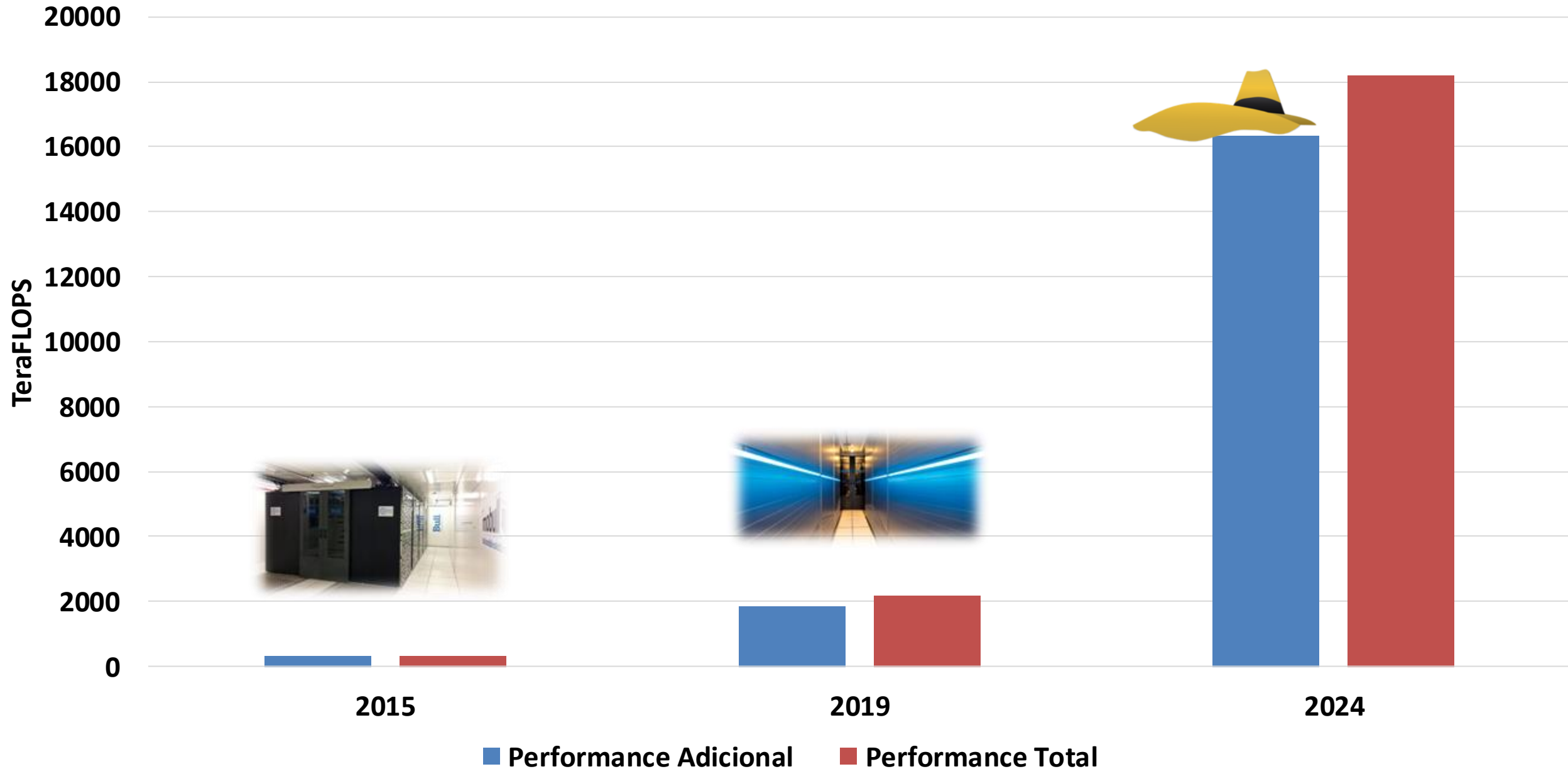
CIÊNCIA, TECNOLOGIA,
INOVAÇÕES E COMUNICAÇÕES



PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

1005

Capacidade Computacional do Laboratório Nacional de Computação Científica



Capacidade Computacional do Centro de HPC do LNCC: Santos Dumont (2015)

Métrica de Comparação	Quantidade
Nós computacionais	558
Processadores(1) por Nó	2
Cores por Processador	12
Cores por Nó	24
Memória por Nó	64GB
Memória por Core	2,6GB
TFLOPS por Nó	0,6373
TFLOPS por Core	0,0265
Rede de Comunicação	IB FDR (\sim 56Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	198
GPUs por Nó	2
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	12GB
Memória por GPU Core	?
TFLOPS por Nó	\sim 2,38 (2xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB FDR (\sim 56Gb/s)

(1) Processador Intel Xeon E5-2695v2 2,4GHz, Ivy Bridge EP

(2) NVIDIA® Tesla® K40, arquitetura NVIDIA Kepler, pico teórico de 1,68 TeraFLOPS por GPU

(3) Fonte: <https://www.top500.org/system/178569/>, <https://www.top500.org/system/178568/>, e <https://www.top500.org/system/178570/>

Capacidade Computacional do Centro de HPC do LNCC: Santos Dumont Expansão (2019)

Métrica de Comparação	Quantidade
Nós computacionais	288
Processadores(1) por Nó	2
Cores por Processador	24
Cores por Nó	48
Memória por Nó	384GB ou 768GB
Memória por Core	8GB ou 16GB
TFLOPS por Nó	1,7718
TFLOPS por Core	0,0369
Rede de Comunicação	IB EDR (100 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	94
GPUs por Nó	4
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	32GB (HBM2)
Memória por GPU Core	?
TFLOPS por Nó	~14,7 (4xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB EDR (100 Gb/s)

(1) Processador Intel Xeon 6252 2,1GHz, Cascade Lake

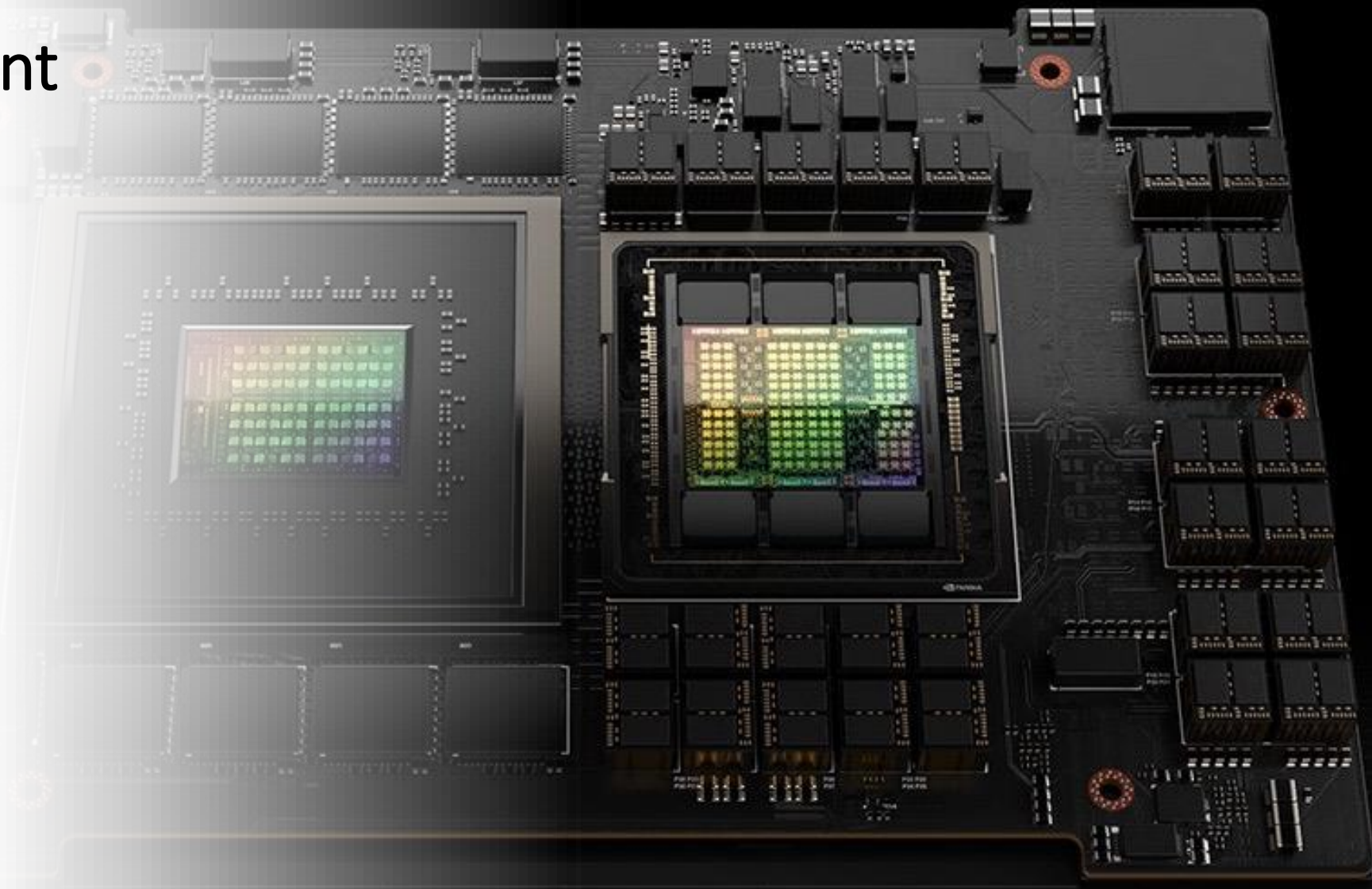
(2) NVIDIA® Tesla® V100 SXM2, arquitetura NVIDIA Volta, pico teórico de 7,8 TeraFLOPS por GPU

(3) Fonte: <https://www.top500.org/system/179704/>

Cluster Acadêmico 2: O Novo Santos Dumont

HETEROGÊNEO
(baseado em GPU)

- **60 nós de CPU:**
 - 2 processadores AMD Genoa 9684X 96c, 2.55 GHz e 1.5 TB DDRAM
- **62 nós de GPU:**
 - 2 procesadores Intel SHR M9468 HBM2 48c, 2 GHz, 1 TB DRAM e 4 Nvidia Hopper HGX H100 80GB HBM3
- **36 nós NVIDIA Grace Hopper:**
 - 1 NVIDIA Max-Pm composto por 4 sockets NVIDIA GH200 Grace Hopper Superchip nodes, 96 GB HBM3 por GPU e 128 GB LPDDR por CPU
- **4 nós computacionais APU AMD nodes:**
 - 4 AMD MI300A
- **4 nós computacionais ARM:**
 - 1 NVIDIA Grace CPU Superchip
- **4 nós computacionais Intel PVC:**
 - 4 Intel Xe PVC Data Center GPU Max 1550 64 GB



Capacidade Computacional do Centro de HPC do LNCC: Novo Santos Dumont (2024)

- **60 nós de CPU:**
 - 2 processadores AMD Genoa 9684X 96c, 2.55 GHz e 1.5 TB DDRAM
- **62 nós de GPU:**
 - 2 procesadores Intel SHR M9468 HBM2 48c, 2 GHz, 1 TB DRAM e 4 Nvidia Hopper HGX H100 80GB HBM3
- **36 nós NVIDIA Grace Hopper:**
 - 1 NVIDIA Max-Pm composto por 4 sockets NVIDIA GH200 Grace Hopper Superchip nodes, 96 GB HBM3 por GPU e 128 GB LPDDR por CPU
- **4 nós computacionais APU AMD nodes:**
 - 4 AMD MI300A
- **4 nós computacionais ARM:**
 - 1 NVIDIA Grace CPU Superchip
- **4 nós computacionais Intel PVC:**
 - 4 Intel Xe PVC Data Center GPU Max 1550 64 GB

Tipo de Nó	HPL (por tipo de nó) (TFLOPs sustentado)	HPL (por nó) (TFLOPs sustentado)
60 CPU	487,12	7,85
62 GPU SPR+H100	9.139,30	147,40 (CPU+GPU)
36 Grace Hopper Superchip	5.553,90	154,27 (CPU+GPU)
4 AMD APU MI300A	804,83	201,21 (CPU+GPU)
4 Grace CPU Superchip	19,79	4,95
4 Intel PVC	340,20	85,05

(1) Processador AMD Genoa 9684X 96c, 2.55 GHz
(2) NVIDIA® Hopper® H100 SXM4, arquitetura NVIDIA Hopper, pico teórico de 34 TeraFLOPS por GPU

Capacidade Computacional do
Centro de HPC do LNCC: Novo Santos Dumont (2024)

Métrica de Comparação	Quantidade
Nós computacionais	60
Processadores(1) por Nó	2
Cores por Processador	96
Cores por Nó	192
Memória por Nó	1,5TB
Memória por Core	8GB
TFLOPS por Nó	7,8548
TFLOPS por Core	0,0409
Rede de Comunicação	IB NDR (200 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	62
GPUs por Nó	4
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	80GB (HBM3)
Memória por GPU Core	?
TFLOPS por Nó	~61,2 (4xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB NDR (200 Gb/s)

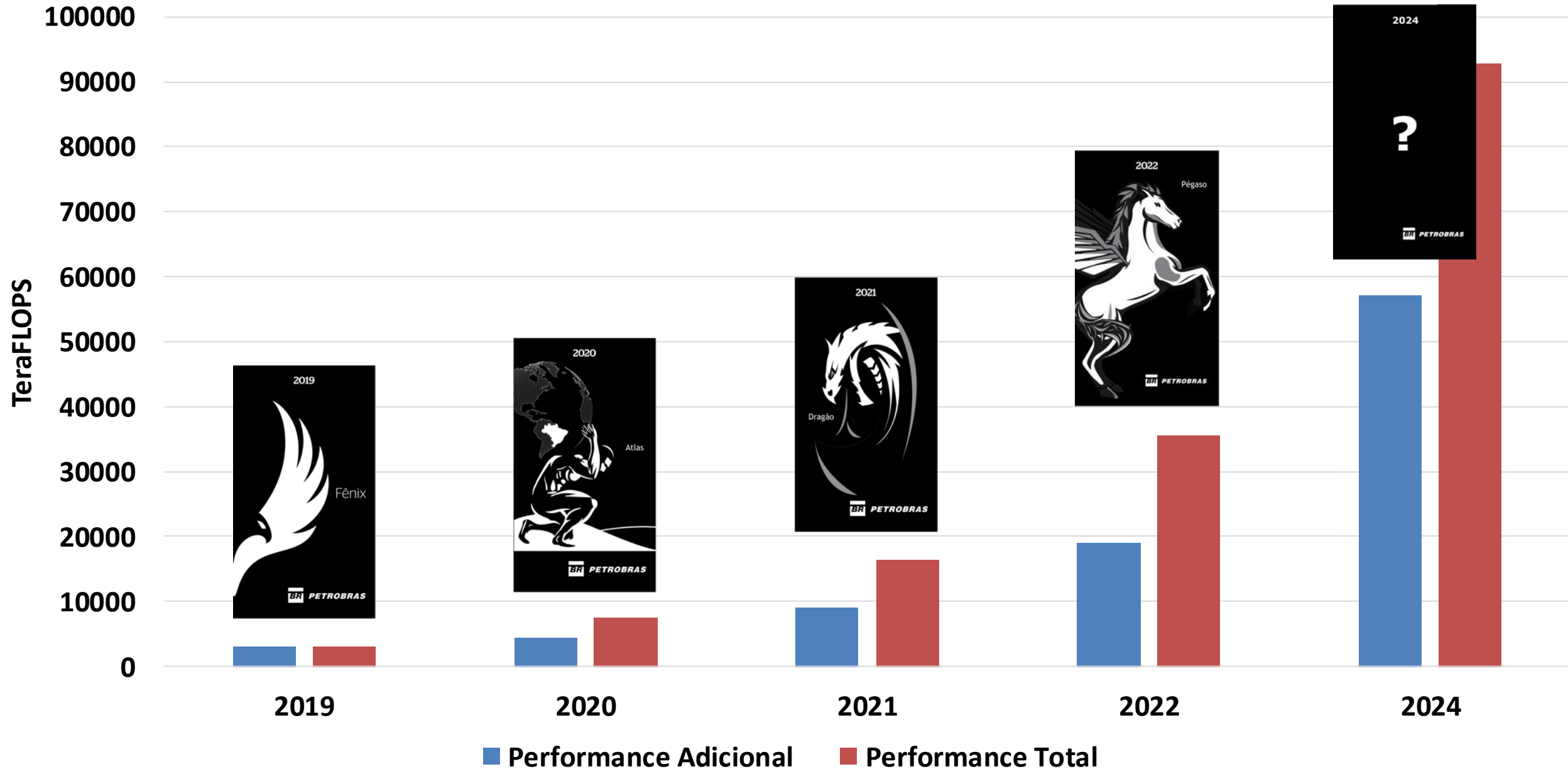
(1) Processador AMD Genoa 9684X, 2.55 GHz


(2) NVIDIA® Hopper® H100 SXM4, arquitetura NVIDIA Hopper, pico teórico de 34 TeraFLOPS por GPU



PETROBRAS

Capacidade Computacional da Petrobras (Imageamento Sísmico GPU)





Fênix



PETROBRAS

Capacidade Computacional da Petrobras, Imageamento Sísmico GPU: Fênix (2019)

Métrica de Comparação	Quantidade
Nós computacionais	360
Processadores(1) por Nó	2
Cores por Processador	4
Cores por Nó	8
Memória por Nó	192GB
Memória por Core	24GB
TFLOPS por Nó	0,2544
TFLOPS por Core	0,0318
Rede de Comunicação	IB EDR (100 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	360
GPUs por Nó	2
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	32GB (HBM2)
Memória por GPU Core	?
TFLOPS por Nó	~8,8 (2xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB EDR (100 Gb/s)

(1) Processador Intel Xeon Skylake-SP 5122, 3.6 GHz

(2) NVIDIA® Tesla® V100, arquitetura NVIDIA Volta, pico teórico de 7 TeraFLOPS por GPU

(3) Fonte: <https://www.top500.org/system/179681/>



Capacidade Computacional da Petrobras, Imageamento Sísmico GPU: Atlas (2020)

Métrica de Comparação	Quantidade
Nós computacionais	136
Processadores(1) por Nó	2
Cores por Processador	18
Cores por Nó	36
Memória por Nó	768GB
Memória por Core	21,3GB
TFLOPS por Nó	1,6447
TFLOPS por Core	0,0457
Rede de Comunicação	IB EDR (100 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	136
GPUs por Nó	8
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	32GB (HBM2)
Memória por GPU Core	?
TFLOPS por Nó	~32 (8xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB EDR (100 Gb/s)

(1) Processador Intel Xeon 6240, 2.6 GHz, Cascade Lake

(2) NVIDIA® Tesla® V100 SXM2, arquitetura NVIDIA Volta, pico teórico de 7,8 TeraFLOPS por GPU

(3) <https://www.top500.org/system/179854/>



Dragão

BR PETROBRAS

Capacidade Computacional da Petrobras, Imageamento Sísmico GPU: Dragão (2021)

Métrica de Comparação	Quantidade
Nós computacionais	272
Processadores(1) por Nó	2
Cores por Processador	26
Cores por Nó	52
Memória por Nó	768GB
Memória por Core	14,8GB
TFLOPS por Nó	1,9192
TFLOPS por Core	0,0369
Rede de Comunicação	IB EDR (100 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	272
GPUs por Nó	8
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	32GB (HBM2)
Memória por GPU Core	?
TFLOPS por Nó	~33 (8xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB EDR (100 Gb/s)

(1) Processador Intel Xeon 6230R, 2.1 GHz, Cascade Lake

(2) NVIDIA® Tesla® V100E SXM2, arquitetura NVIDIA Volta, pico teórico de 7,8 TeraFLOPS por GPU

(3) <https://www.top500.org/system/179941/>



Pégaso

PETROBRAS

Capacidade Computacional da Petrobras, Imageamento Sísmico GPU: Pégaso (2022)

Métrica de Comparação	Quantidade
Nós computacionais	252
Processadores(1) por Nó	2
Cores por Processador	32
Cores por Nó	64
Memória por Nó	2048GB
Memória por Core	32GB
TFLOPS por Nó	2,6689
TFLOPS por Core	0,0417
Rede de Comunicação	IB HDR (200 Gb/s)

Métrica de Comparação	Quantidade
Nós computacionais	252
GPUs por Nó	8
Cores por GPU	?
Cores GPU por Nó	?
Memória por GPU	80GB
Memória por GPU Core	?
TFLOPS por Nó	~ 75 (8xGPU)
TFLOPS por GPU Core	?
Rede de Comunicação	IB HDR (200 Gb/s)

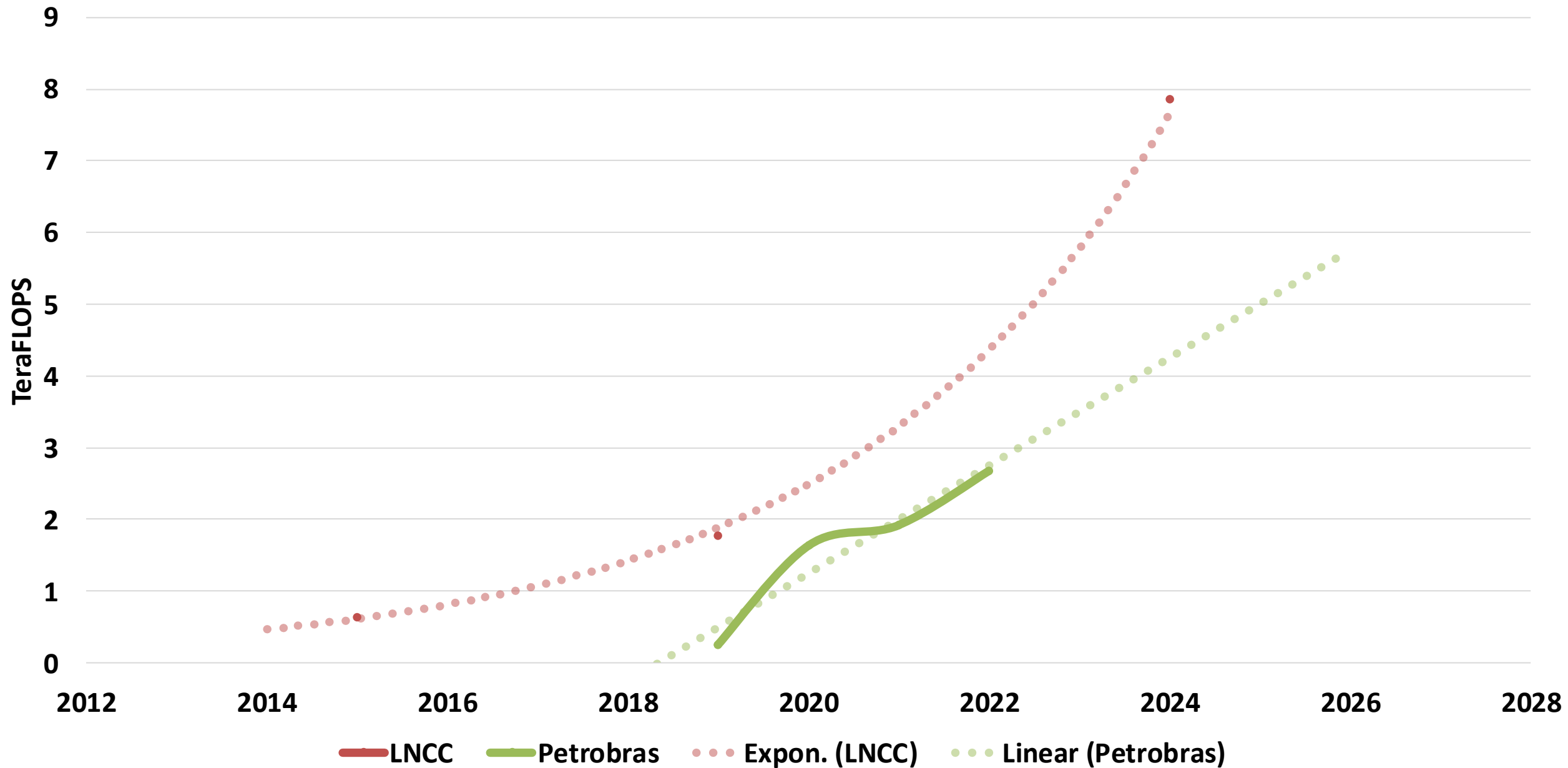
(1) Processador AMD Milan 7513, 2.6 GHz

(2) NVIDIA A100 SXM4, arquitetura NVIDIA Ampere, pico teórico de 19,5 TeraFLOPS por GPU (Tensor Core)

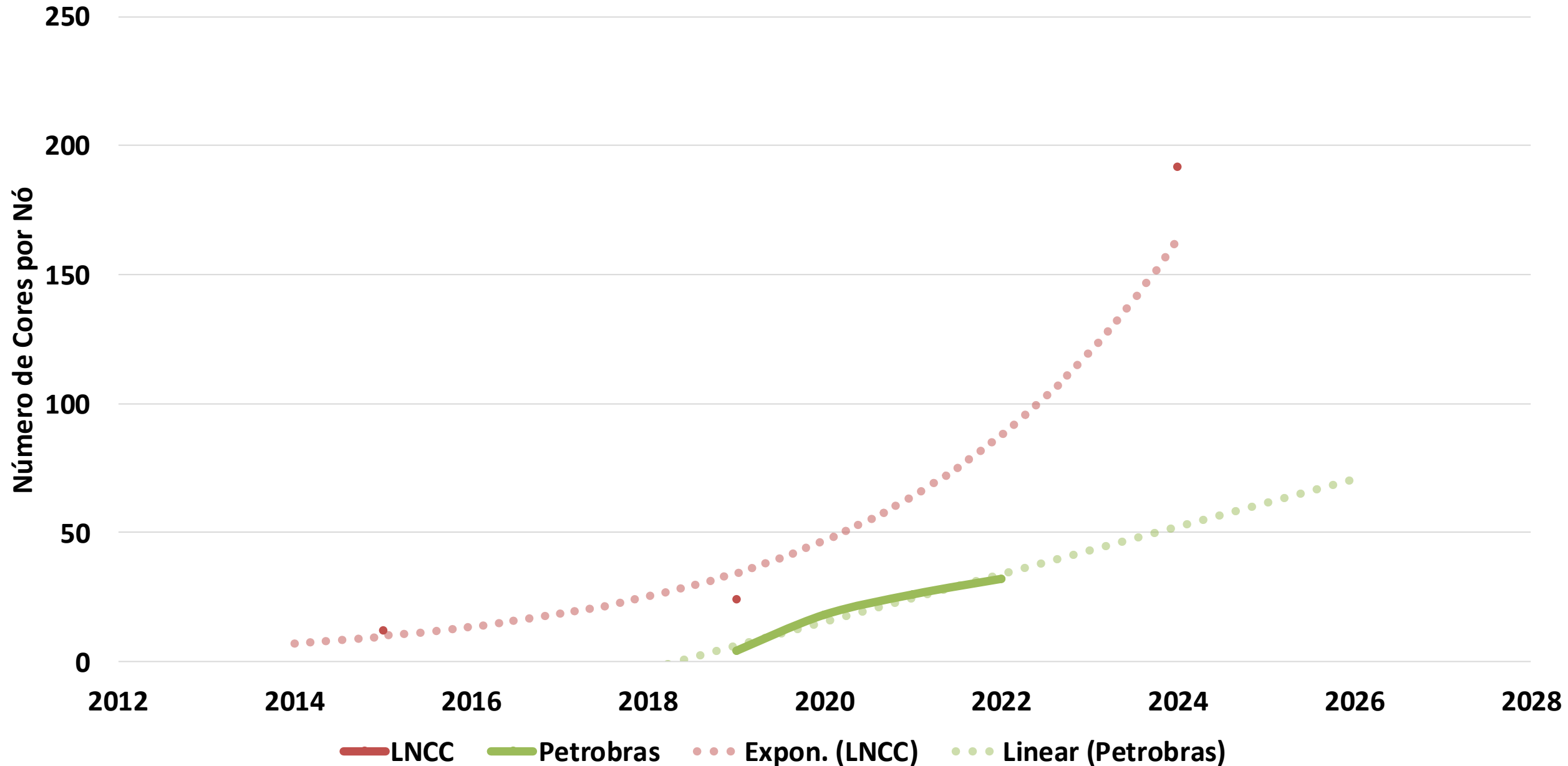
(3) <https://www.top500.org/system/180127/>

Análises e Pontos de Discussão

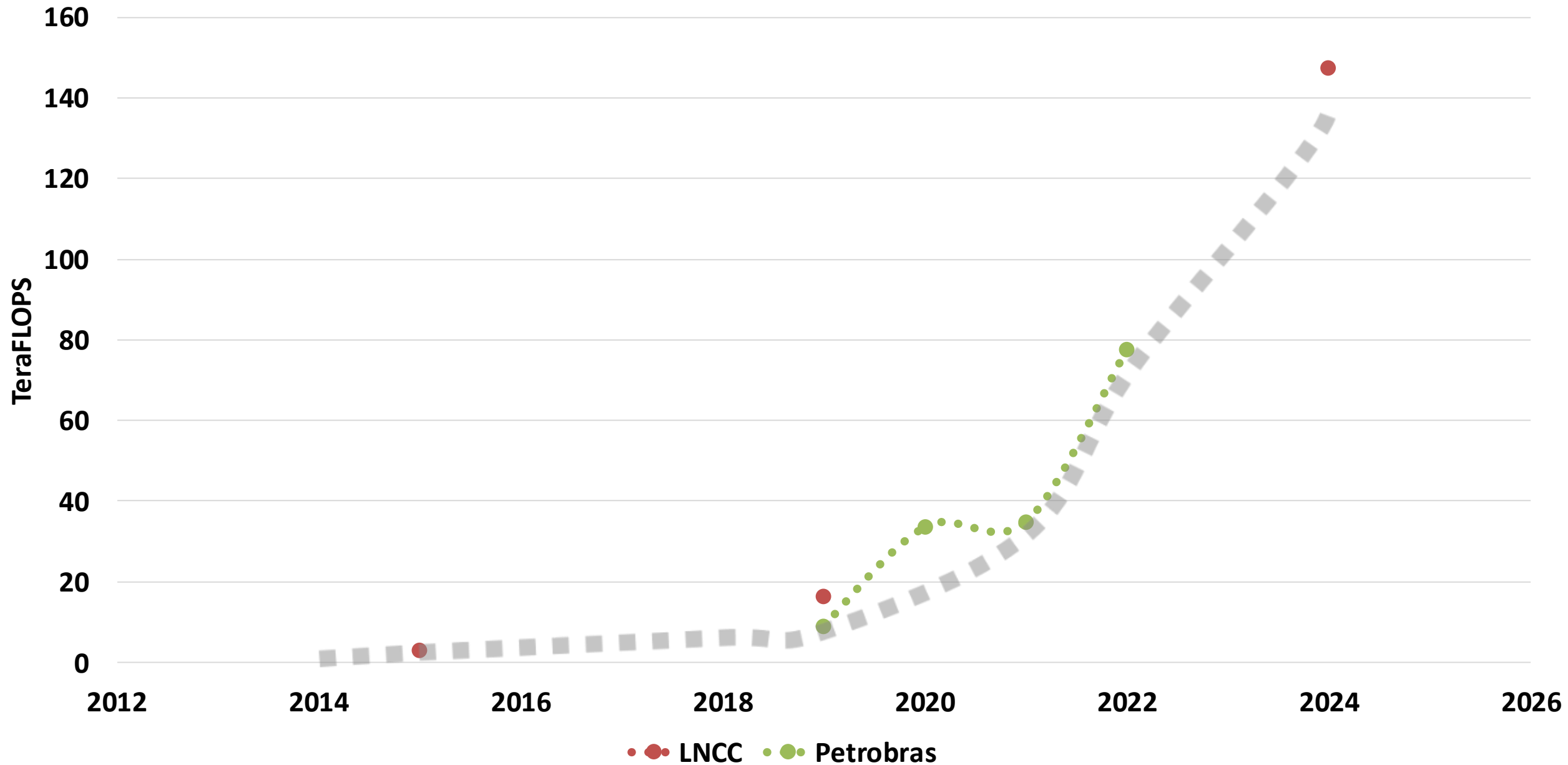
Evolução da Capacidade de Processamento em CPU por Nó de Processamento



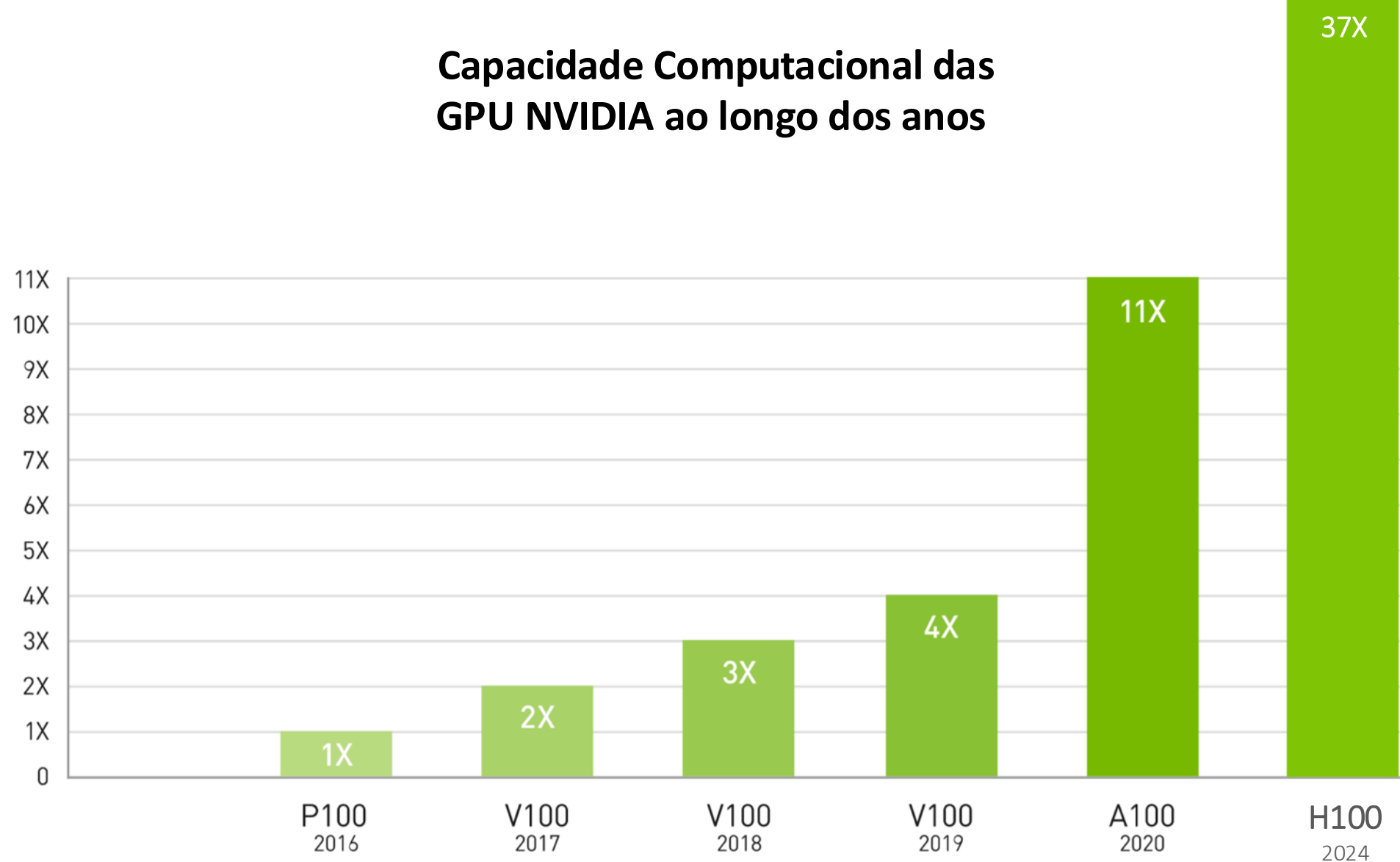
Evolução do Número de Cores de CPU por Nó de Processamento



Evolução da Capacidade de Processamento em CPU+GPU por Nó de Processamento



Capacidade Computacional das GPU NVIDIA ao longo dos anos



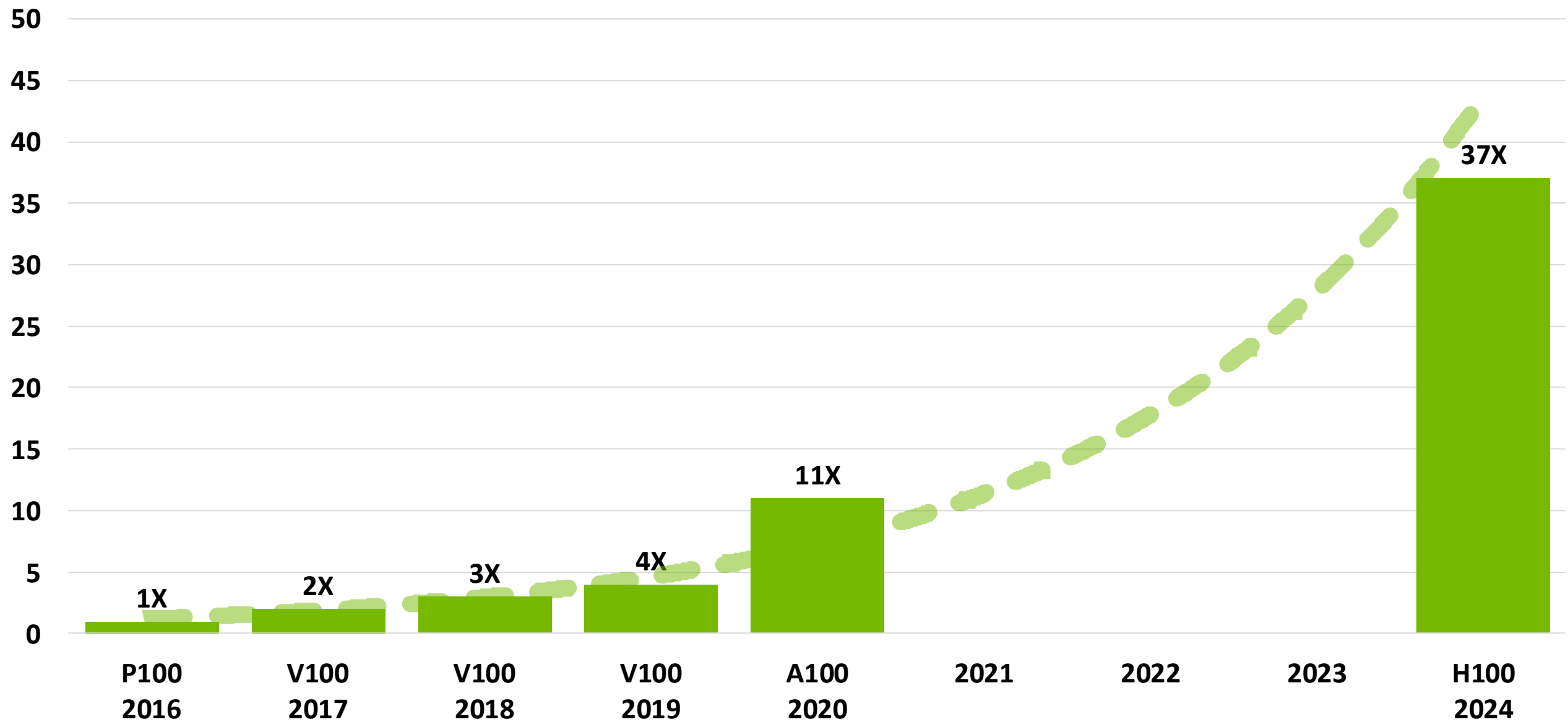
Throughput - Relative Performance

(1) Fonte: <https://www.nvidia.com/pt-br/data-center/a100/>

(2) Média geométrica de acelerações de aplicação vs. P100: Aplicação de referência: Âmbar [PME-Cellulose_NVE], Chroma [szscl21_24_128], GROMACS [ADH Dodec], MILC [Apex Médio], NAMD [stmv_nve_cuda], PyTorch (BERT-Large Fine Tuner), Quantum Espresso [AUSURF112-jR], Random Forest FP32 [make_blobs (160000 x 64: 10)], TensorFlow [ResNet-50], VASP 6 [Si Huge] | Nó de GPU com CPUs de soquete duplo com 4x NVIDIA P100, V100, ou GPUs A100.

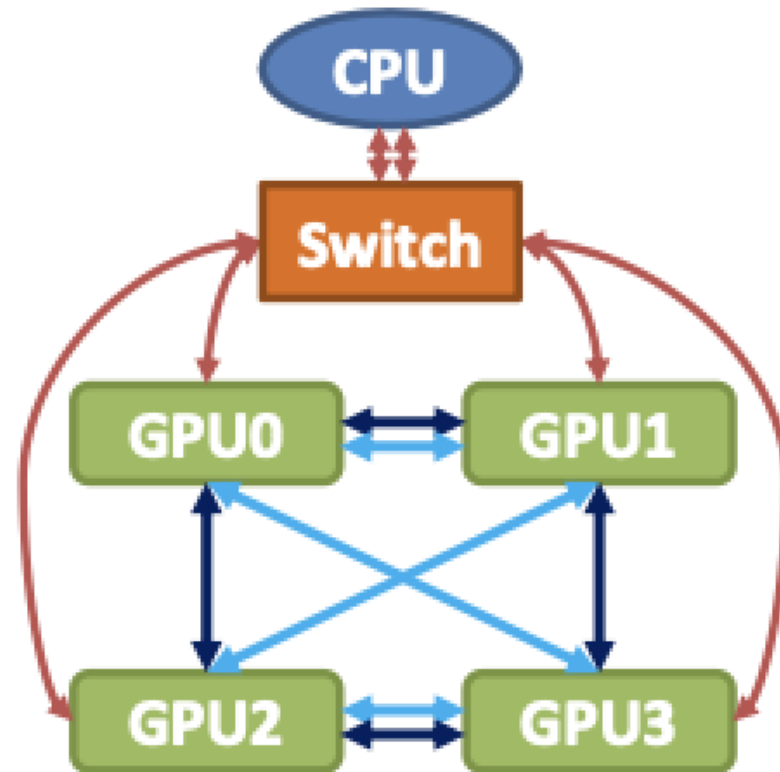
(3) A100 (tensor cores) performance is 19,7 TFlops, H100 (tensor cores) performance is 67 Tflops (3,4 times higher)

Capacidade Computacional das
GPU NVIDIA ao longo dos anos
(performance relativa)

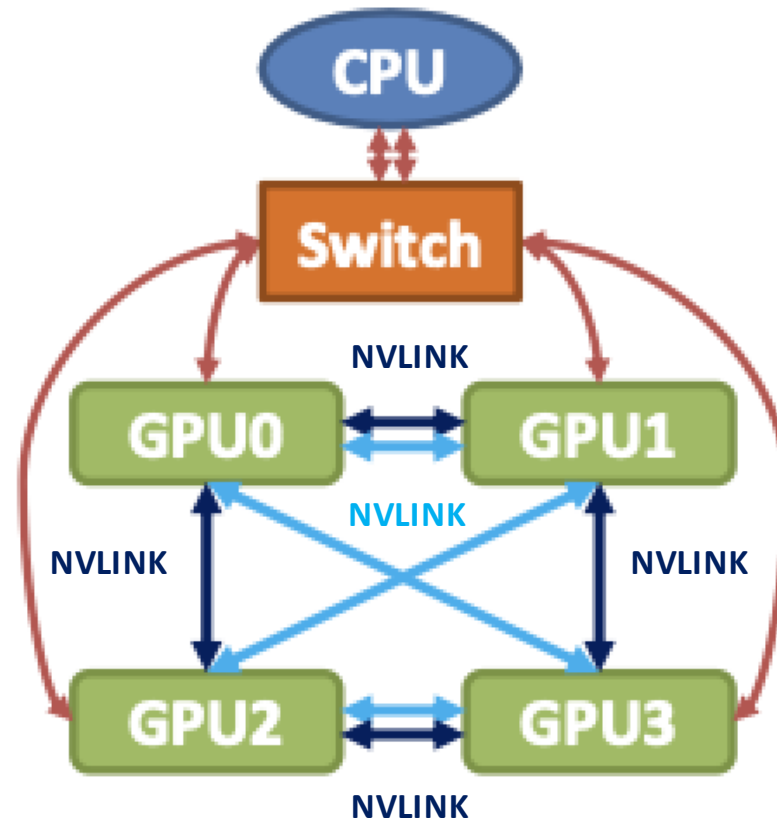


Sistemas Multi-GPU

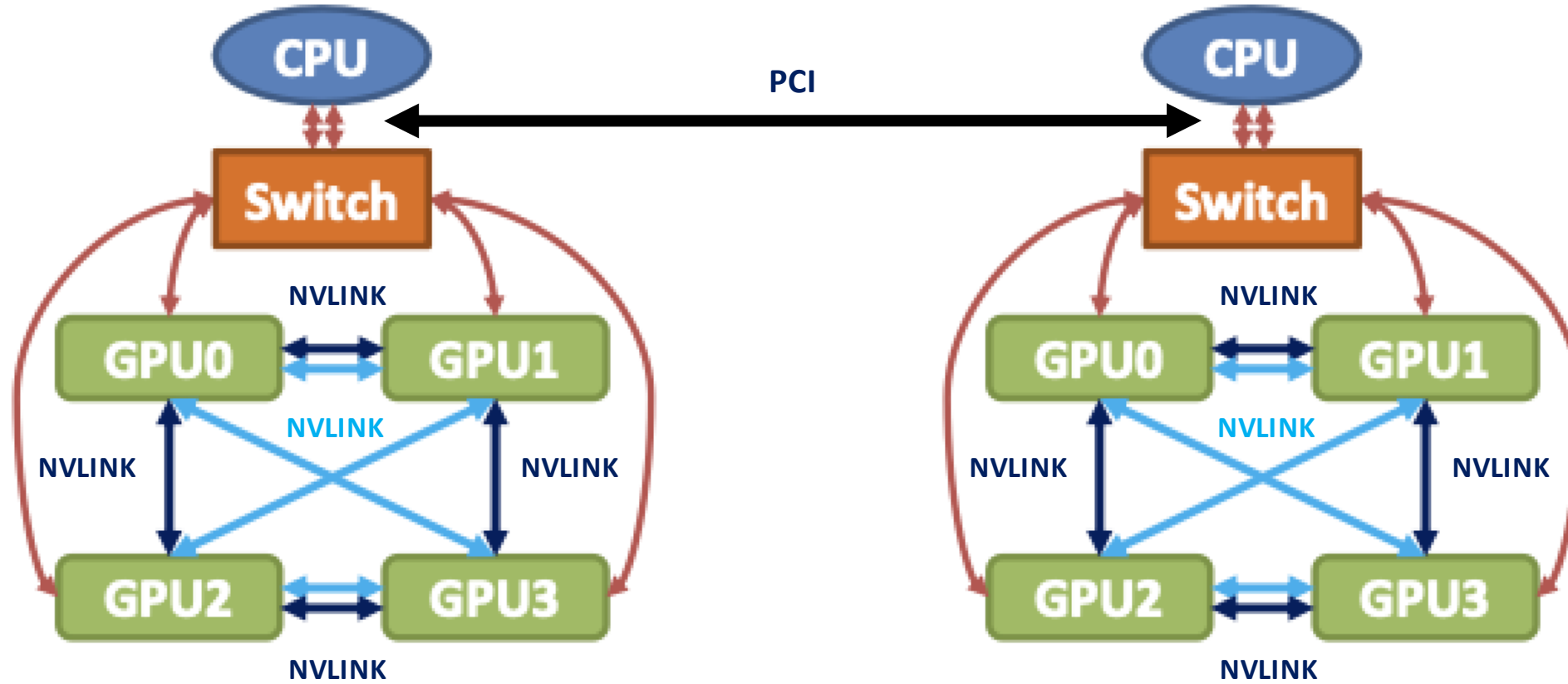
Multi-GPU Nodes



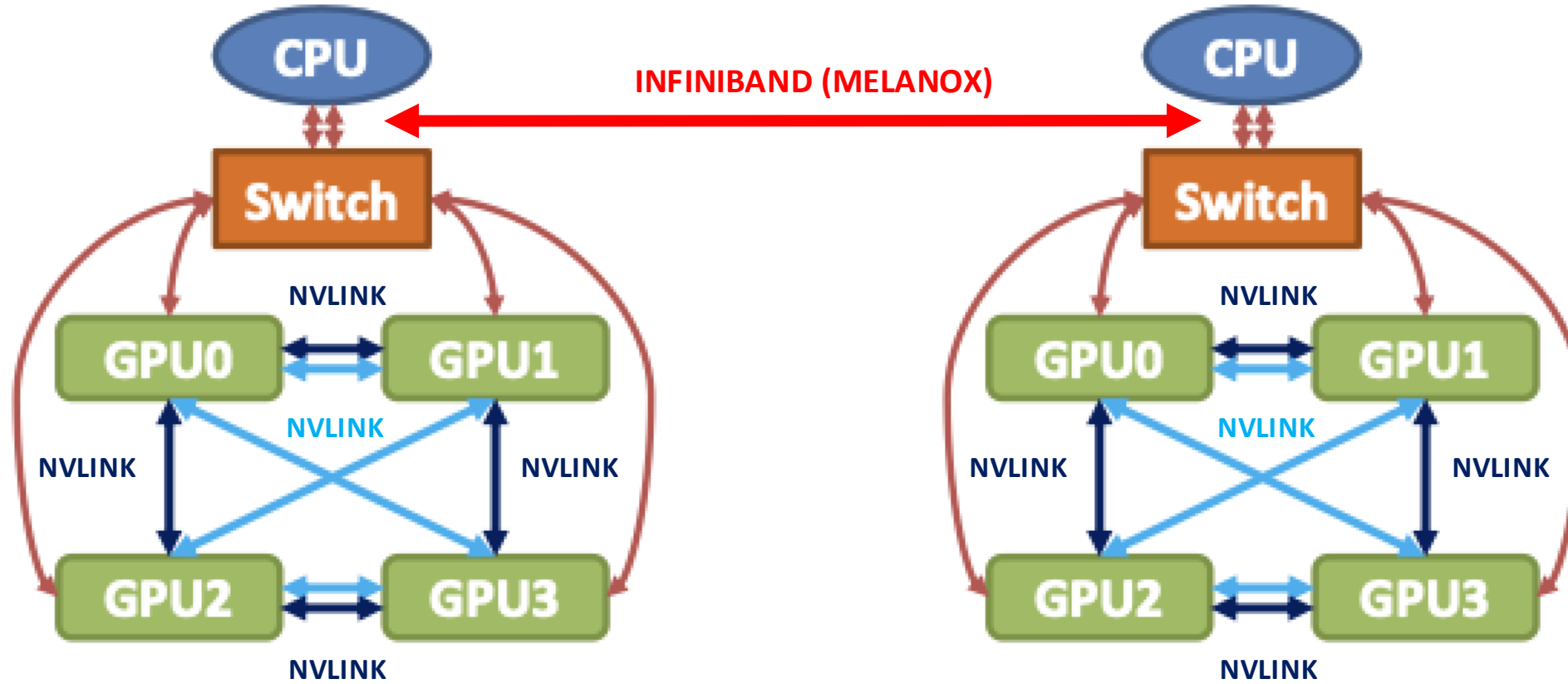
Multi-GPU Nodes



Multi-GPU Nodes

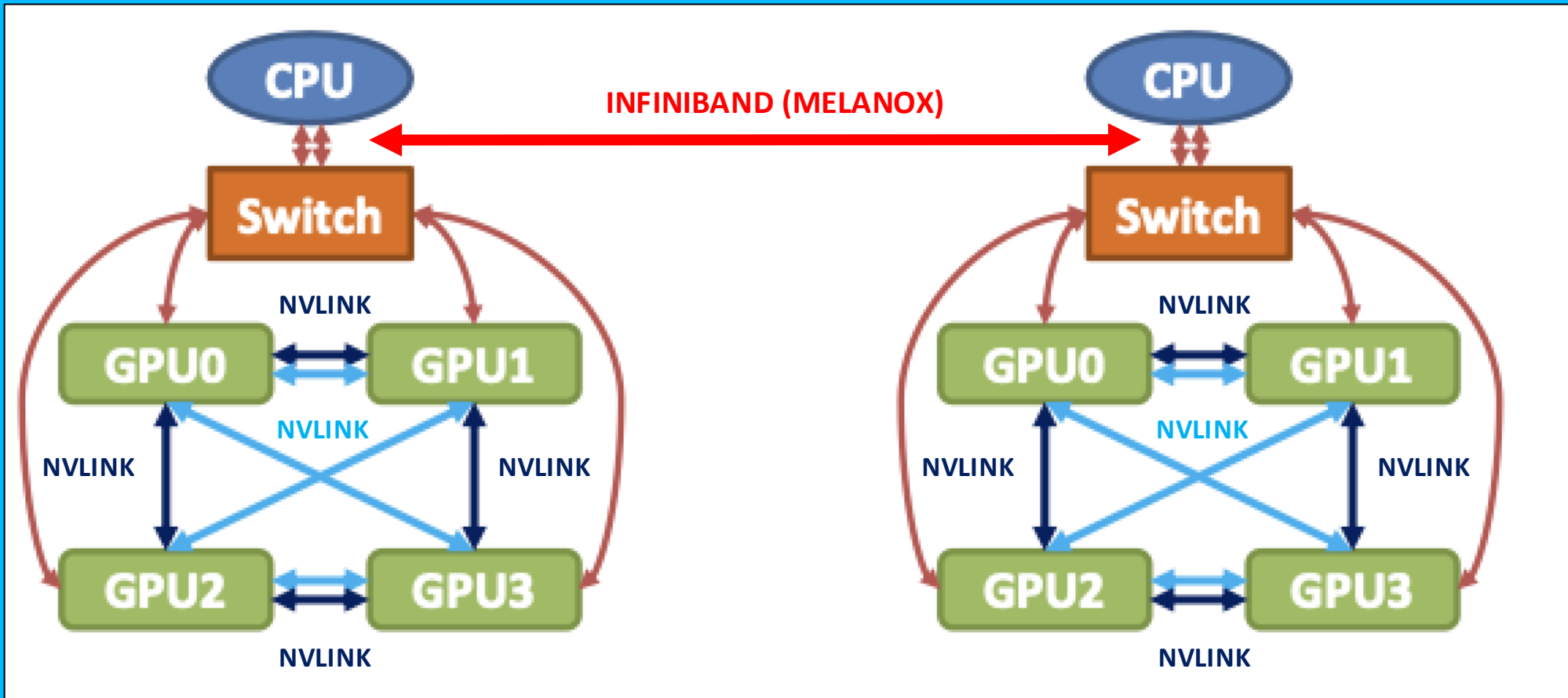


Multi-GPU Nodes



Multi-GPU Nodes

UCX (Unified Communication X Library)



Multi-GPU Environments

OPENACC

CUDA

OPENMP

NCCL

NVLINK

UCX

RDMA

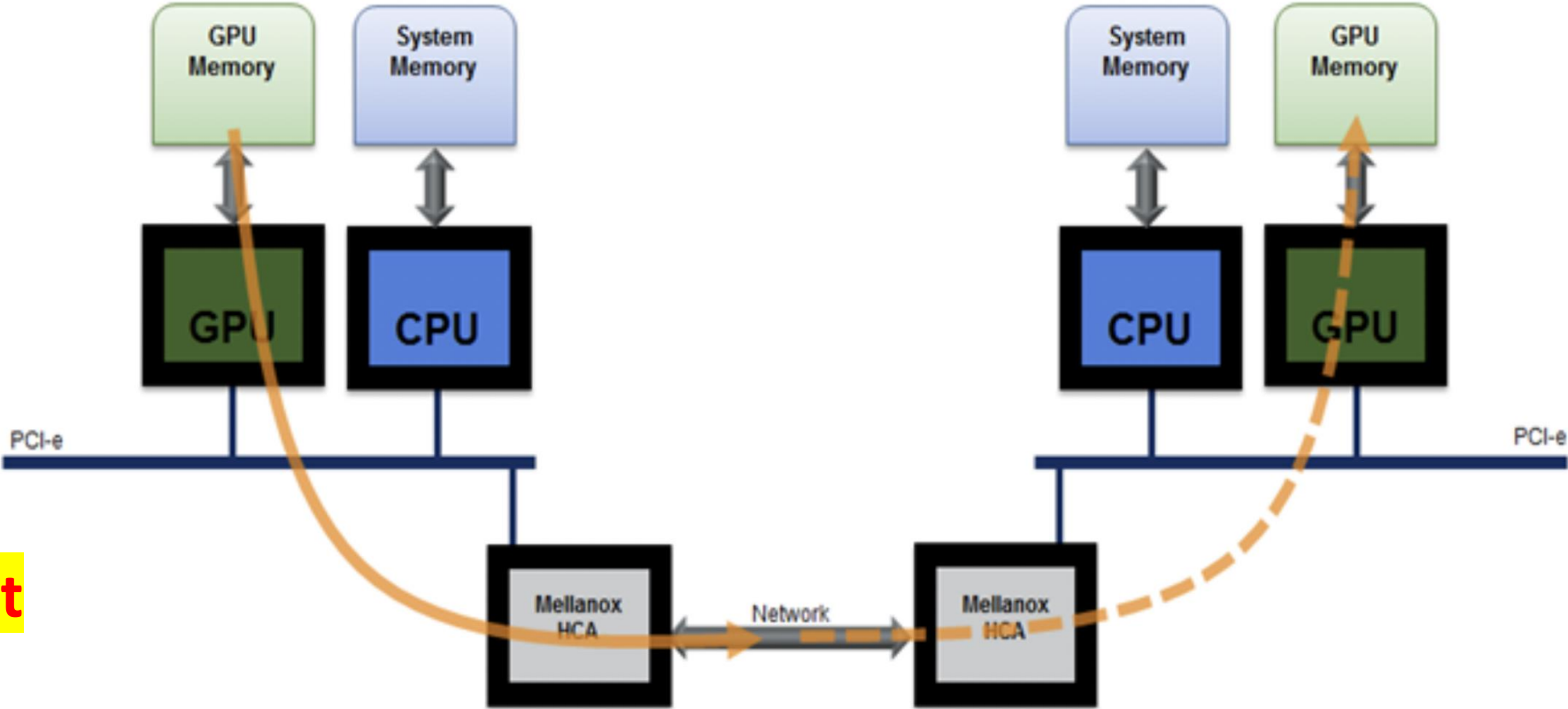
GPUDirect

MPI+CUDA

INFINIBAND (MELLANOX)

CUDA-aware

NVSHMEM



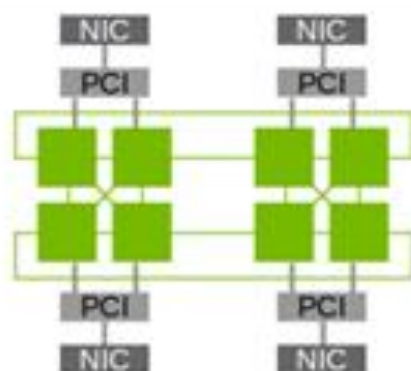
NCCL

**NVIDIA Collective Communication
Library**

Topology detection

Build graph with all GPUs, NICs, CPUs, PCI switches, NVLink, NVSwitch.

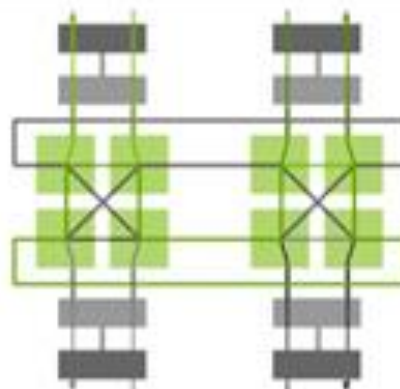
Topology injection for VMs.



Graph search

Extensive search to find optimal set of rings or trees.

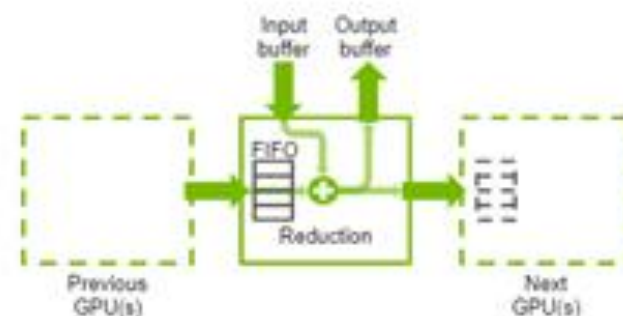
Performance prediction of each algorithm and auto-tuning.



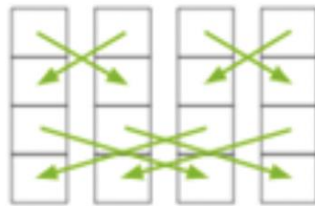
CUDA kernels

Optimized reductions and copies for a minimal SM usage.

CPU threads for network communication.



sendrecv



gather



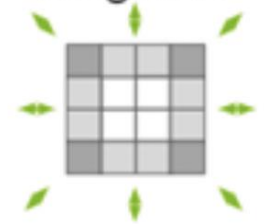
scatter



alltoall



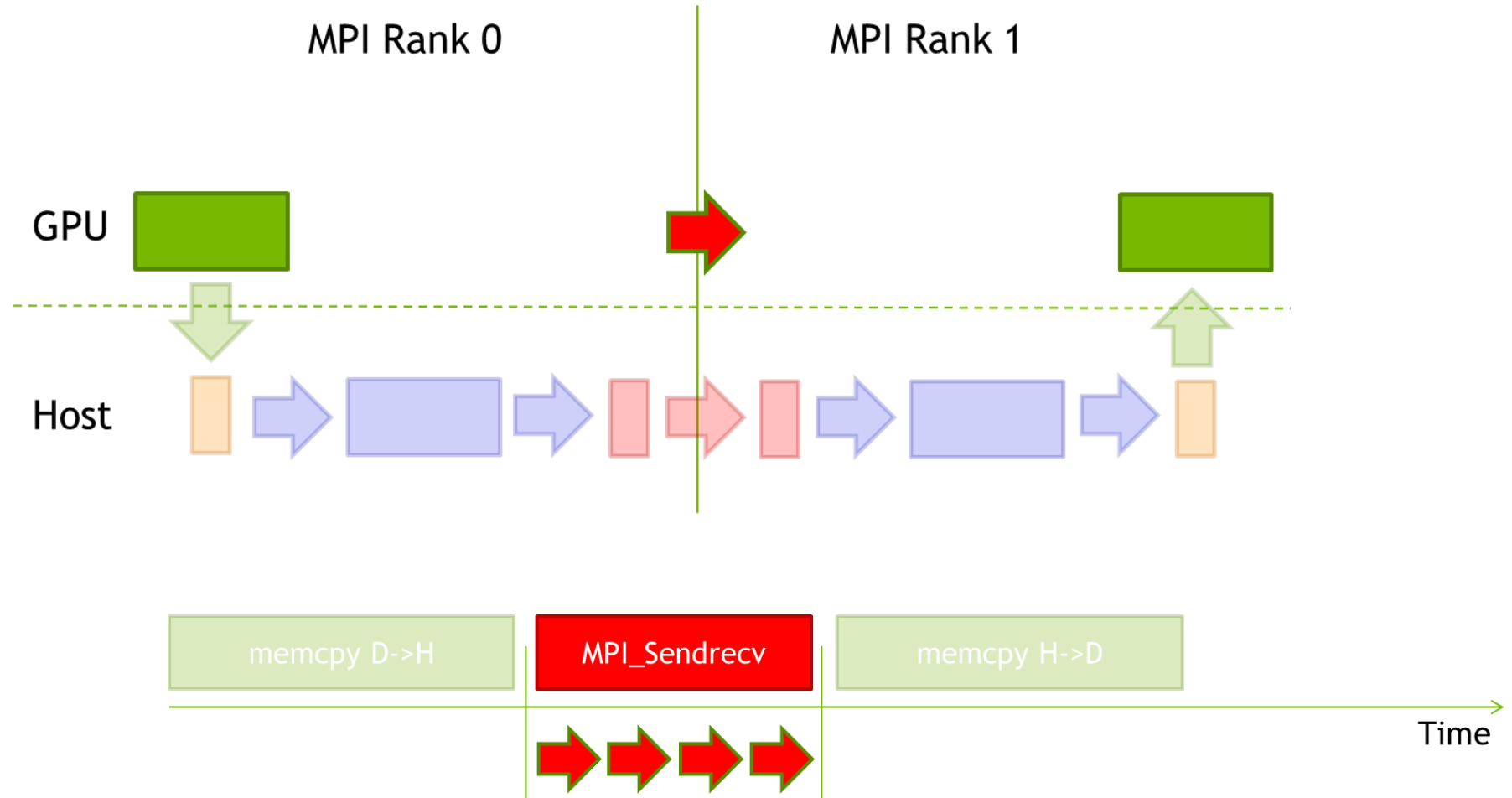
neighbor



CUDA-AWARE MPI

CUDA + MPI at GPUs

CUDAWARE-MPI

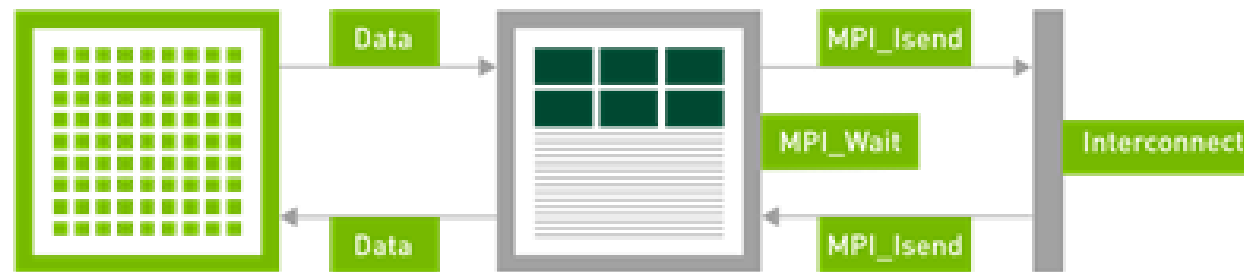


NVSHMEM

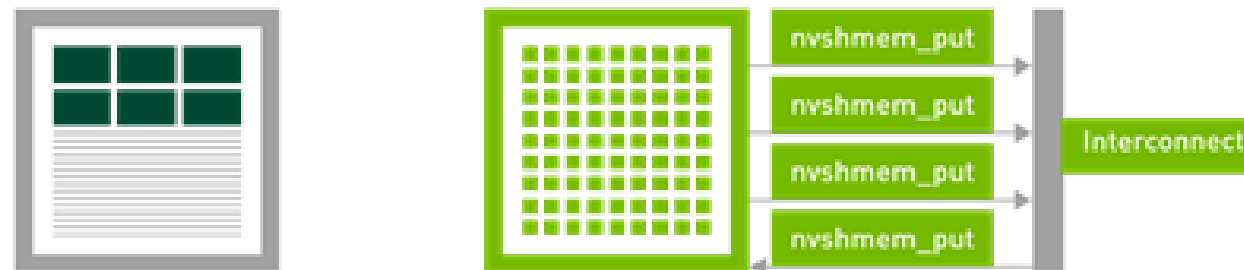
NVIDIA OpenSHMEM

NVSHMEM

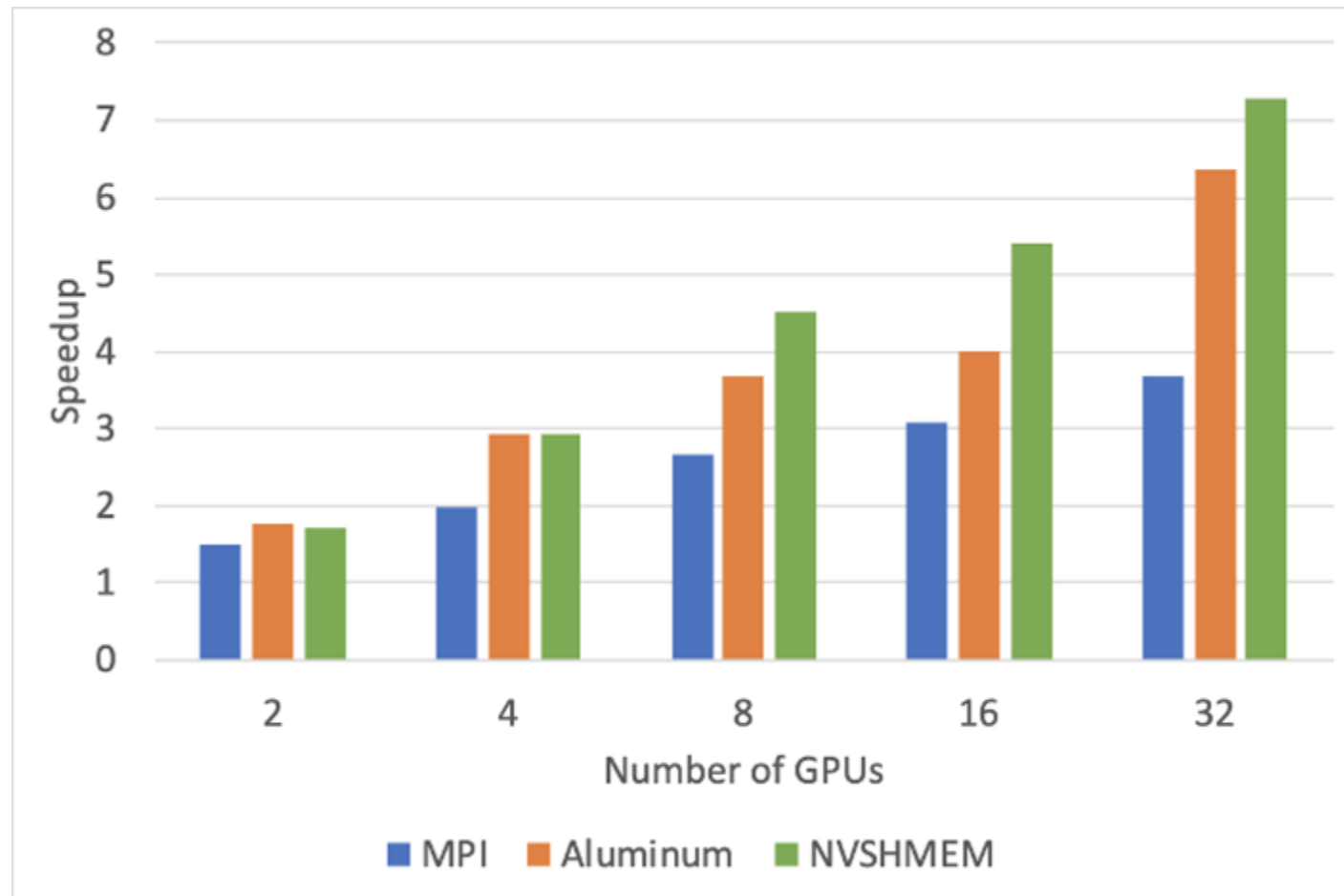
MPI



NVSHMEM



NVSHMEM



Conclusão do Pontos Discutidos

Como estimar quantos nós de CPU (exclusivamente) e quantos híbridos (CPU+GPU)?

- A. É uma tendência todos os nós terem GPU(s)?
- B. Como (e quais aplicações) usar **eficientemente** os nós híbridos?

As aplicações conseguem escalar eficientemente dentro de um mesmo nó?

- A. Caso de nós de CPU com centenas de cores.**
- B. Caso de nós com GPU, inclusive com múltiplas GPUs (e centenas de cores CPU).**
- C. Memória compartilhada ou memória distribuída? Híbrido?**
- D. Deveríamos pensar em compartilhar (múltiplos jobs) os nós de CPU e híbridos?**

HPC as a Solution

