Hardware Heterogeneity at NIST

DEREK JUBA

2023-12-07

NIST Clusters

Raritan

- ~1000 heterogenous nodes, mostly Intel CPU, mostly no GPU

Enki

- 13 nodes, IBM Power9 CPU, Nvidia V100 GPU

Nisaba

- 4 nodes (Nvidia DGX), Intel Xeon CPU, Nvidia V100 GPU

Many others...

Information Systems Group Cluster

Nodes
- 21 Heterogeneous

CPUs
- Intel Xeon, AMD Ryzen, AMD EPYC

GPUs
- None, TITAN RTX, GeForce 3090, V100, P100, A10, A100 (PCIE-40GB, PCIE-80GB, SXM4-80GB), H100

RAM
- 128 GiB to 2 TiB

Points of Heterogeneity

CPU Architecture: x86, Power ISA

- Extensions: FMA, SSE, AVX, etc.

GPU "Architecture": Nvidia, AMD, Intel

- Compute Capability

RAM Size

Exotic Hardware

- Tensor Processing Units, Quantum Processing Units, FPGAs, Xeon Phi, etc.

CUDA Backward and Forward Compatibility

CUDA SDK version(s)	Tesla	Fermi	Kepler (early)	Kepler (late)	Maxwell	Pascal	Volta	Turing	Ampere	Ada Lovelace	Hopper
1.0	1.0 – 1.1										
1.1	1.0 – 1.1+x										
2.0	1.0 – 1.1+x										
2.1 - 2.3.1	1.0 – 1.3										
3.0 - 3.1	1.0	2.0									
3.2	1.0	2.1									
4.0 - 4.2	1.0	2.1									
5.0 - 5.5	1.0			3.5							
6.0	1.0			3.5							
6.5	1.1				5.x						
7.0 - 7.5		2.0			5.x						
8.0		2.0				6.x					
9.0 - 9.2			3.0				7.0				
10.0 - 10.2			3.0					7.5			
11.0				3.5					8.0		
11.1 - 11.4				3.5					8.6		
11.5 - 11.7.1				3.5					8.7		
11.8				3.5							9.0
12.0 - 12.3					5.0						9.0

Source: Wikipedia

NIST Issues with CUDA Forward Compatibility

Received container which included pytorch

- Initially ran on V100 GPU
- Later, could not run on A100 GPU

Why?

NIST Issues with CUDA Forward Compatibility

CUDA Compilation

- NVCC compiler translates CUDA to PTX
- CUDA driver translates PTX to executable binary

CUDA applications are meant to ship with PTX

- CUDA driver compiles PTX for user's hardware at runtime

Pytorch shipped with only executable binaries...

- Just update Pytorch!
- Not in a container...

Thank You

ANY MENTION OF COMMERCIAL PRODUCTS IS FOR INFORMATION ONLY; IT DOES NOT IMPLY RECOMMENDATION OR ENDORSEMENT BY NIST.