

AMD ROCm

計算機結構報告



111210543資工二 余宸安

目錄

01.簡介

02.核心概念

03.CUDA vs ROCm

04.ROCm的組成

05.ROCm原理

簡介

MD ROCm (**Radeon Open Compute Platform**) 是 AMD 開發的開源高效能計算平台，專門用於支持異構計算和加速科學運算、深度學習、機器學習等應用。

核心概念

ROCm 是一個針對 GPU 加速運算的軟體平台，旨在提供高效能和靈活性，適合多種異構計算工作負載。它支持開發者在 AMD GPU 和其他支援 OpenCL 的硬體上進行運算。

CUDA vs ROCm

特性	CUDA	ROCm
開發者	NVIDIA	AMD
開源性	專有, 非開源	完全開源
支援的硬體	只支援NVIDIA GPU	主要支援 AMD GPU, 也有一定的 NVIDIA 支援
開發接口	CUDA C/C++、cuDNN、cuBLAS 等	HIP、MIOpen、rocBLAS 等
操作系統	Linux、Windows	主要是 Linux
框架支持	TensorFlow、PyTorch、MXNet 等	TensorFlow、PyTorch 等
性能	通常在 NVIDIA GPU 上表現最佳	在 AMD GPU 上表現強勁, 性能不斷提升

ROCm 的組成

ROCm Runtime: GPU 和 CPU 資源管理、異構計算支持

HIP: 跨平台編程接口, 支持 CUDA 轉換為 ROCm

MIOpen: 深度學習加速庫, 類似於 cuDNN

ROCm Libraries: 高效能計算和深度學習的數學運算庫

ROCm Compiler and Tools: 編譯工具鏈(HCC、clang)、性能分析、調試工具

Framework Support: 支援 TensorFlow、PyTorch 等深度學習框架

ROCd 和 ROCprof: 調試工具和性能分析工具

ROCm 原理

1. 異構計算架構 (Heterogeneous Computing Architecture)

ROCm 基於異構計算架構, 這意味著它能夠在 CPU 和 GPU 之間進行協同運算, 並充分發揮不同處理器的優勢。

2. 編程模型: HIP (Heterogeneous-computing Interface for Portability)

ROCm 提供了一個跨平台編程接口 HIP, 使開發者能夠編寫兼容 AMD 和 NVIDIA GPU 的程式碼。HIP 是 ROCm 的核心編程模型, 旨在幫助開發者將 CUDA 程式碼轉換為可以在 AMD GPU 上運行的程式碼。

3. GPU 資源管理和調度

ROCm 通過 ROCr (Radeon Open Compute Runtime) 來管理 GPU 資源。ROCr 是 ROCm 的運行時層, 負責資源的分配、調度以及異構計算的協同工作。

4. 高效能數學庫

ROCm 提供了多個針對高效能計算和深度學習的數學庫, 這些庫可以幫助開發者加速常見的數學運算, 如線性代數運算、傅里葉變換等。

ROCm 原理

5. 高效能計算 (HPC) 支持

ROCm 支援各種高效能計算應用，包括科學計算、數值模擬、物理模擬等。ROCm 的設計使其能夠有效地支持大規模的計算集群，並且提供分布式計算支持。

6. 性能分析和調試工具

ROCm 提供了一系列性能分析和調試工具，幫助開發者進行應用優化和錯誤排查。

參考資料

英伟达CUDA vs AMD ROCm生态圈差距研究 :https://www.sohu.com/a/753731068_121880839

Nvidia CUDA vs AMD ROCm vs Intel oneAPI AI與HPC軟體堆疊比較:<https://n.yam.com/Article/20240626975442>

What is ROCm?:<https://youtu.be/CnIKvgdJU9Q?si=gLNrJHfA-mJwiRRo>

ROCm平台及HIP介绍: <https://blog.csdn.net/chongbin007/article/details/124043701>