



# Taipei-1 introduction

## - 「Taipei-1」算力申請(CFP) 說明會

Apr 19, 2024

# Taipei-1 introduction

## Co-innovation computing platform

- Taipei-1 是一座 NVIDIA 在台灣投資建置的超級電腦，具備支持人工智慧與數位孿生運算需求的運算資源環境，將投入提供 25% 算力支持產官學研及新創研發使用所需運算資源的基礎設備。
- 人工智慧運算環境提供 NVIDIA 各式 AI 開發環境與套件，例如支持大語言模型與生成式 AI 開發環境 NVIDIA NeMo Framework (詳細說明: <https://developer.nvidia.com/generative-ai>)，例如支持醫療與生命科學的各項 AI 開發環境 NVIDIA Clara (詳細說明: <https://developer.nvidia.com/industries/healthcare>)，完整的軟體清單可參考: <https://catalog.ngc.nvidia.com/containers>
- 數位孿生運算環境提供 NVIDIA Omniverse (<https://www.nvidia.com/zh-tw/omniverse/solutions/digital-twins/>) 運算需求，建構 Omniverse Farm 環境執行大量 Rendering 運算需求 (詳細說明: <https://docs.omniverse.nvidia.com/farm/latest/index.html>)



# Taipei-1 25% 主機環境說明

- **NVIDIA DGX SuperPOD**

- 16台DGX H100，整體共計128個H100 GPU。
- 每台DGX H100主機內具備8個NVIDIA H100 SXM GPU透過最新一代的NVLink以及NVSwitch技術進行整合，任2顆H100彼此之間具備900GB/s的雙向溝通頻寬。
- DGX H100主機與主機之間透過8條InfiniBand技術搭配NVIDIA Quantum-2 InfiniBand Switch 提供400Gb/s高頻寬與低延遲高速網路溝通頻寬。
- 主要運用在AI運算，生成式AI運算與NVIDIA最新的前瞻技術開發。

- **NVIDIA OVX SuperPOD**

- 8台OVX 2.0 主機，整體共計64個L40 GPU。
- 每台主機內具備8個NVIDIA L40 PCIe GPU。
- 透過NVIDIA高速網路架構進行整合。
- 主要運用在Digital Twins數位孿生與NVIDIA最新的前瞻技術開發。

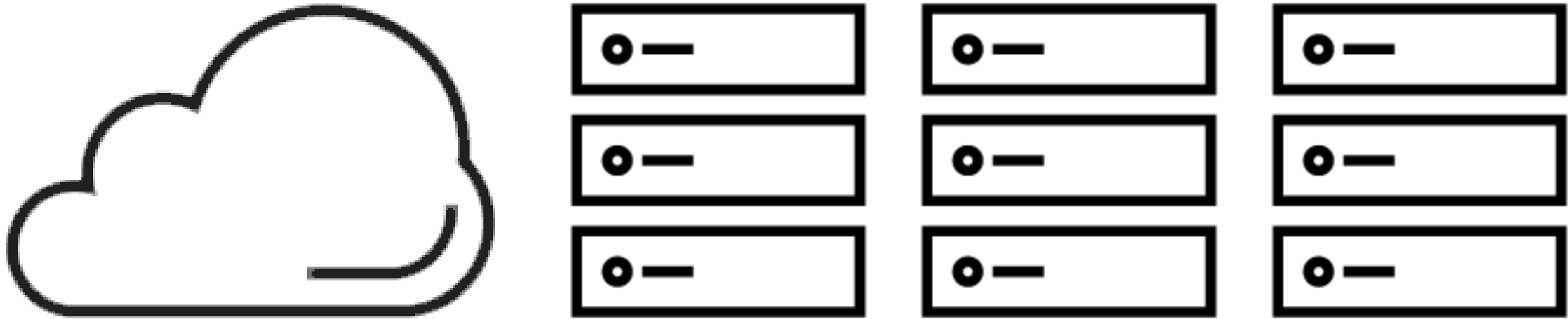
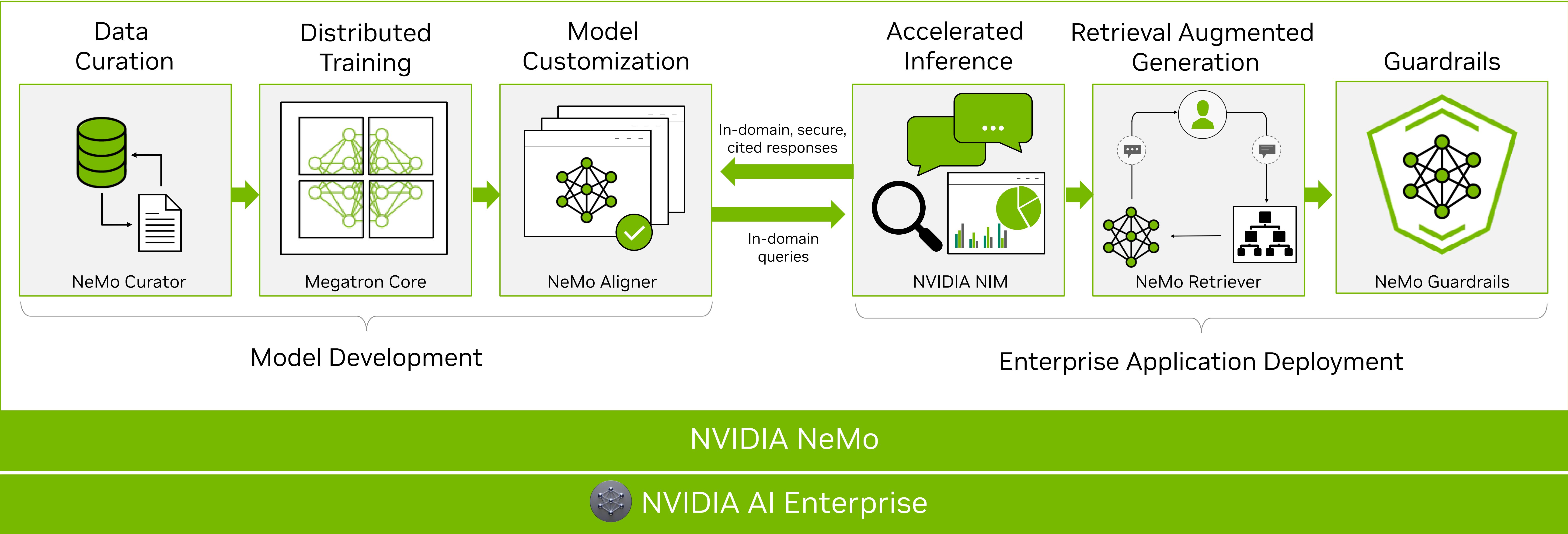
# 適合專案類型

Taipei-1 co-innovation computing platform

- 生成式AI ( **Generative AI** ) 相關應用
- 大型語言模型
- 物理資訊神經網路
- 量子計算模擬
- 藥物發現
- 醫療保健
- 數位孿生

# Generative AI Applications - Large Language Model

Build, customize and deploy generative AI models with NVIDIA NeMo



Support Falcon, Llama2, T5, Mistral 7B, Mixtral 8\*7B

User Guide: <https://docs.nvidia.com/nemo-framework/user-guide/latest/overview.html>

# Use Cases for OVX

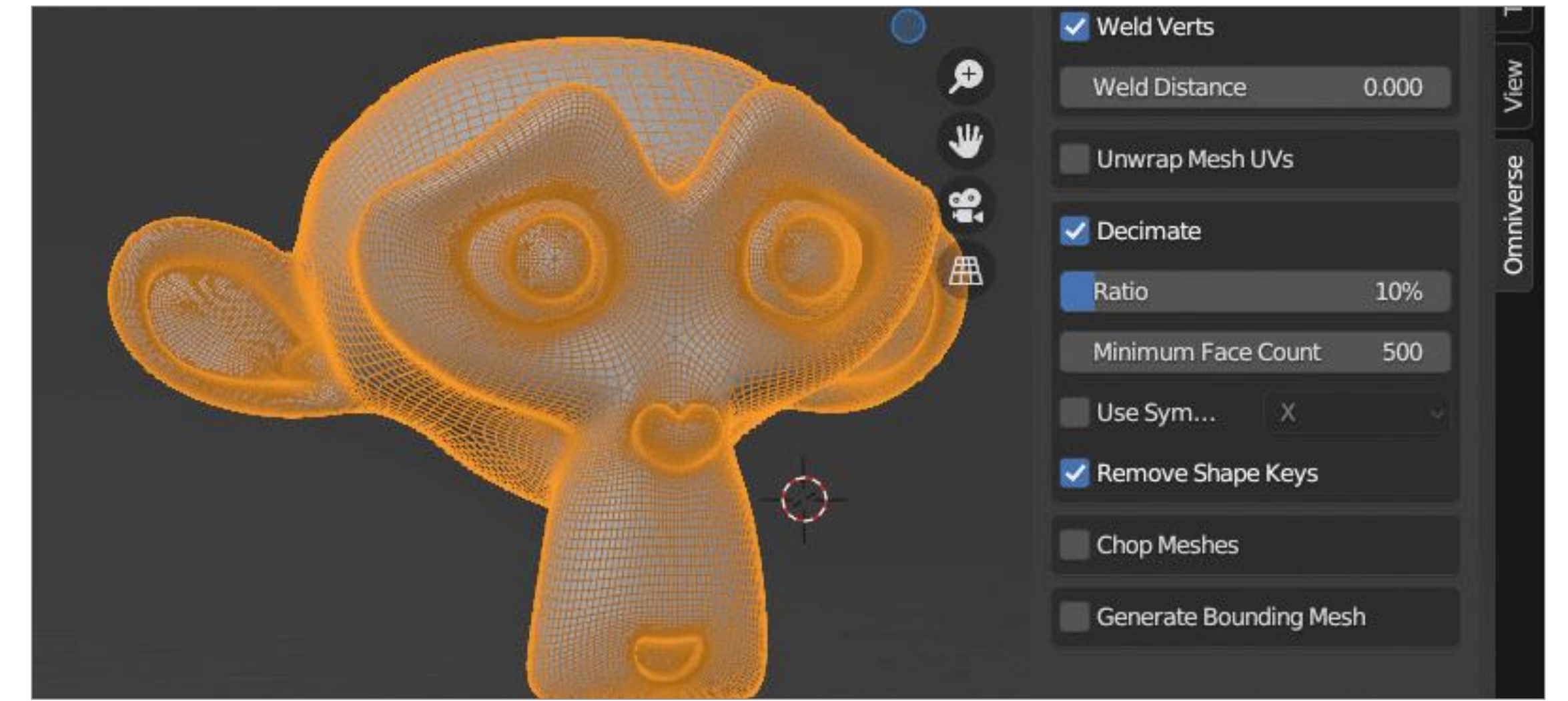
Omniverse Farm



Video Rendering



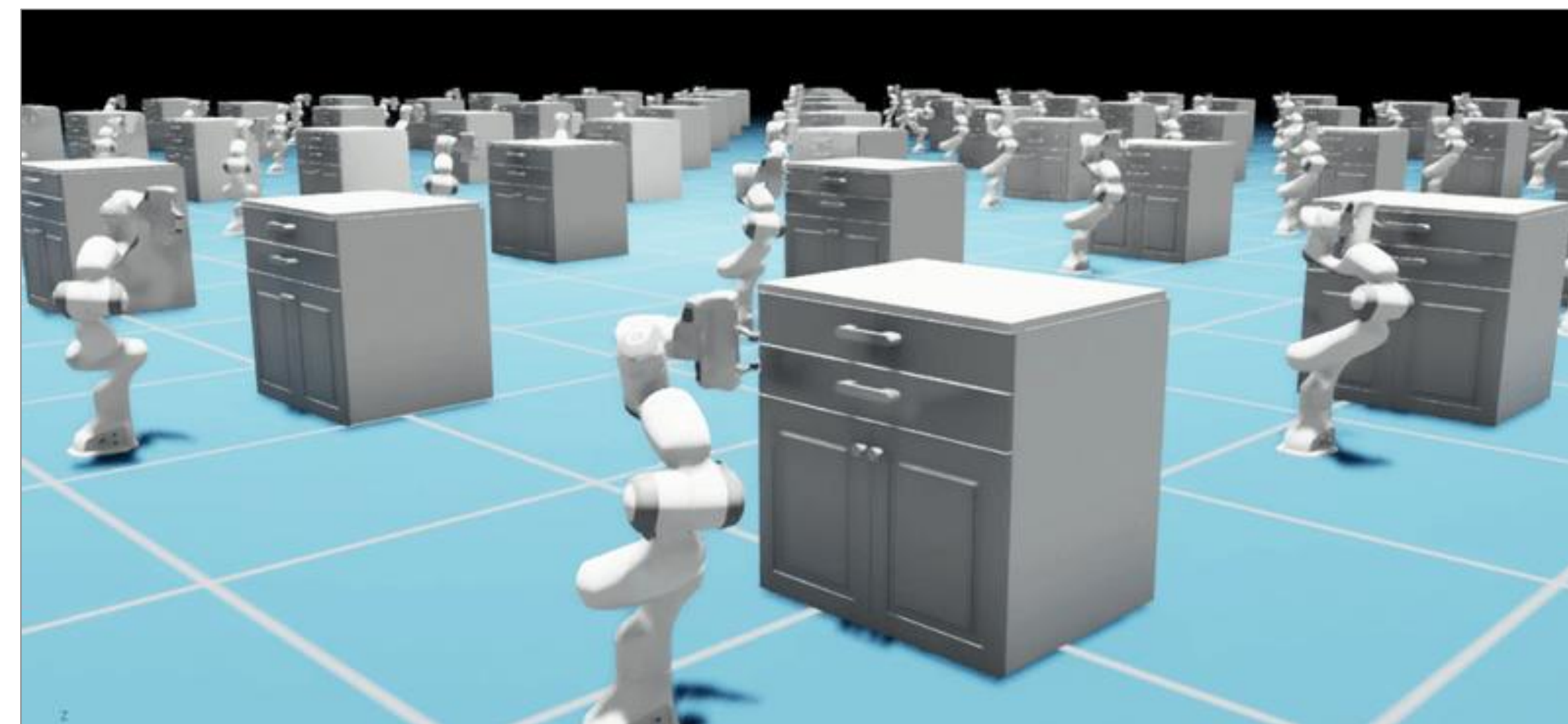
Issac-Sim PhysX Simulation



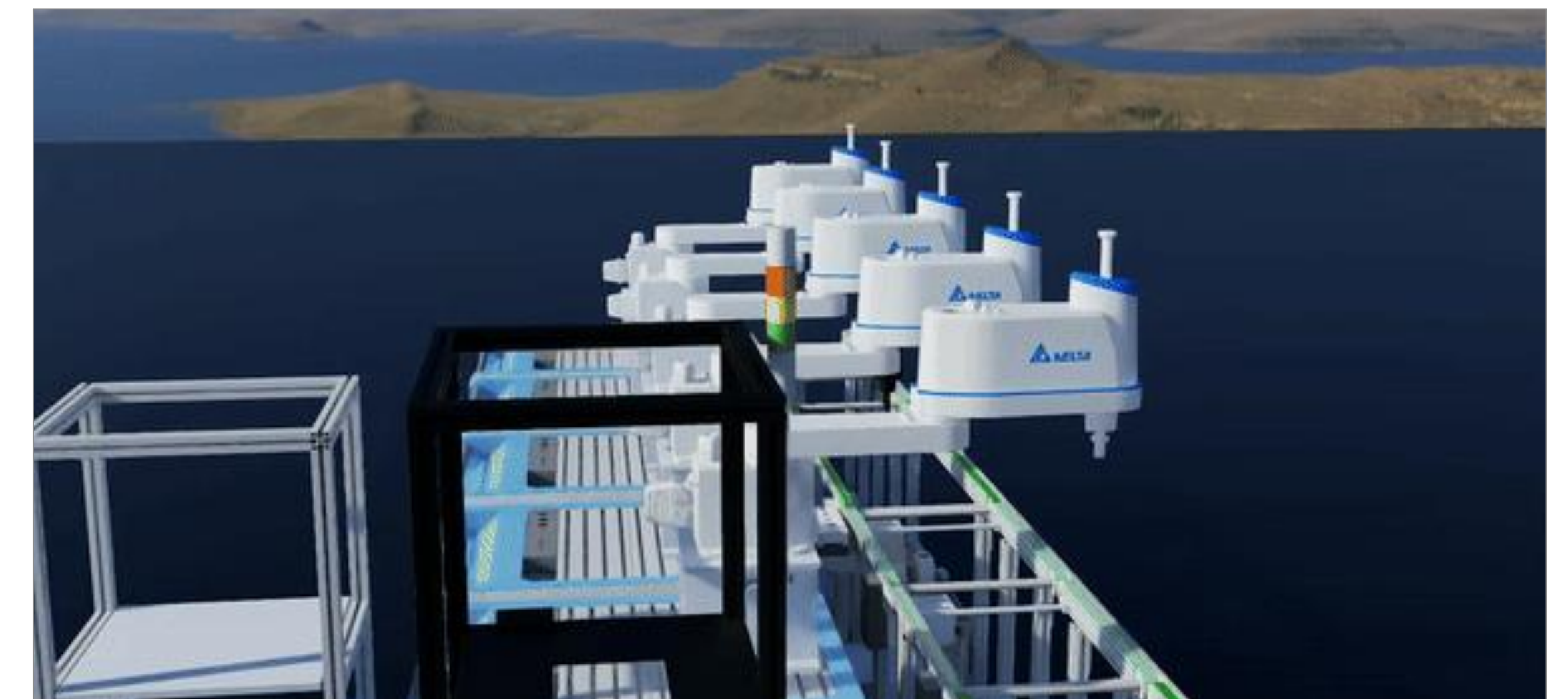
Scene Optimizer



Neural Radiance Field (NeRF)



RL Training



Synthetic Data Generation



**Thanks**