# Reaching the Edge of the Edge: Image Analysis in Space

Robert Bayer, Julian Priest, Pınar Tözün

DEEM @ SIGMOD

9th June 2024

RAD
rad.itu.dk

IT UNIVERSITY OF COPENHAGEN
itu.dk

DASYA
Data-Intensive Systems and Applications
dasya.itu.dk

novo nordisk fonden

# DISCO: Danish Student CubeSat Organization

- Collaboration across Danish universities
- Use-case: Build a CubeSat satellite for observation of land masses (especially snow, ice) in the Arctic

# Small Satellites

## Benefits

$$
⇩
$

Reduced cost

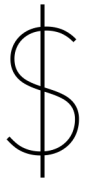
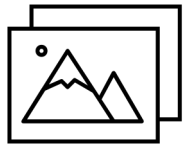
Shrinking
+
Standardization

## Compromises

Reduced
Power
Generation

# Network on Small Satellites

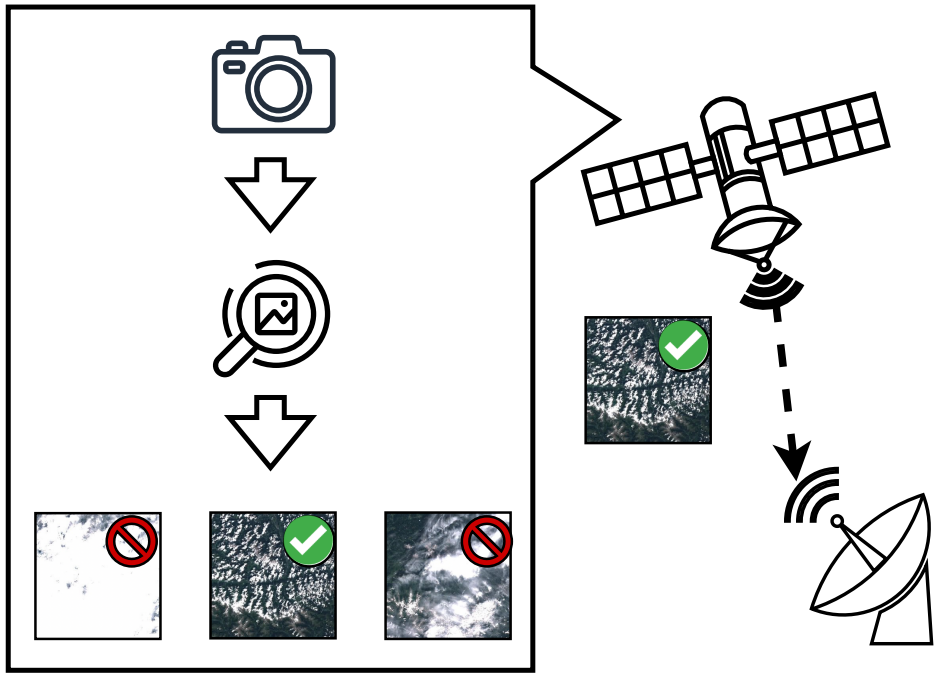


| | | | | |
|---|---|---|---|---|
| UHF | Slow | Low-power | Low | 0.05 / day* |
| S-Band | Fast | High-power | High | 49.1 / day* |

* Based on 4512x4512px image

# Machine Learning Onboard Satellites



Our goal: Determine the right edge device to deploy on the satellite for this task.

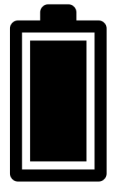
# Requirements

MAX 49.1 images transferred / day!

Real-time imaging
4.42 s latency

Arctic region
71.74 s latency

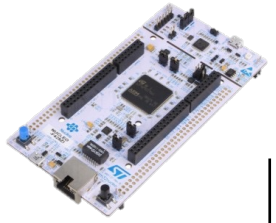
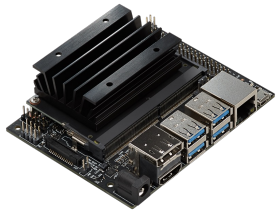
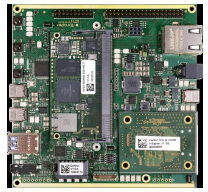
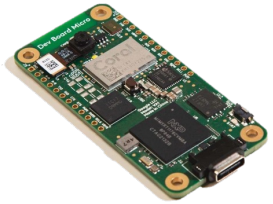
Greenland
270 s latency
320 images captured/day

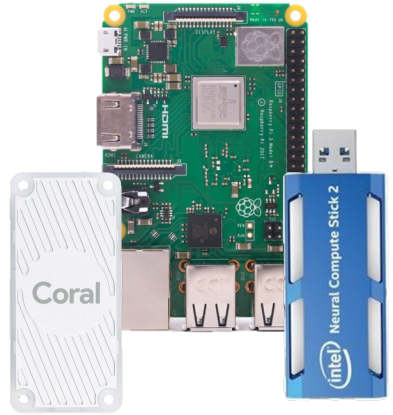


<5W
*MAX*

<2W
*AVG*

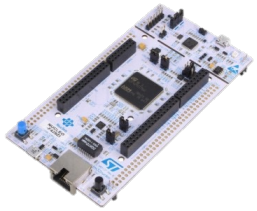
# Devices under Test

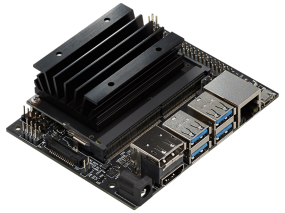





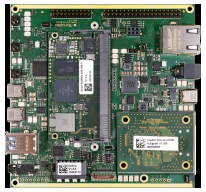



| | ARM Cortex-M7 | Jetson Nano | Toradex Verdin iMX8MP | Coral Dev Board Micro | Coral Dev Board Mini | RPi 3 + Coral USB accel. | RPi 3 + Neural Compute Stick 2 |
|---|---|---|---|---|---|---|---|
| **CPU** | ARM Cortex-M7 @ 300MHz | ARM Cortex-A57 @ 1.43 GHz | ARM Cortex-A53 @ 1.8 GHz | ARM Cortex-M7 @ 800 MHz | ARM Cortex-A35 @ 1.5 GHz | BCM2837 @ 1.2 GHz | |
| **RAM** | 384KB | 4GB | 4GB | 64MB | 2GB | 1GB | |
| **Accelerator** | N/A | 128-core Maxwell GPU | ARM NPU | Coral Edge TPU | | | Intel Neural Compute Stick 2 |

# Experimental Setup





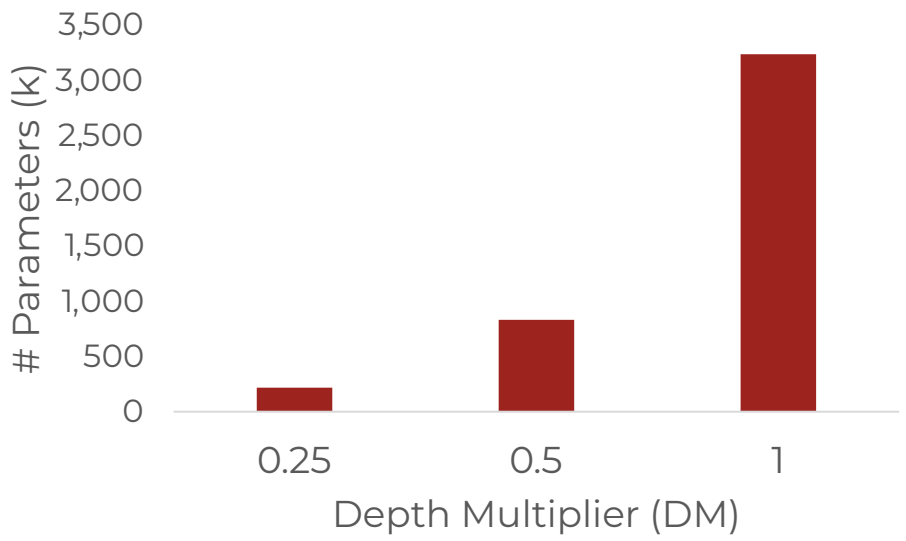




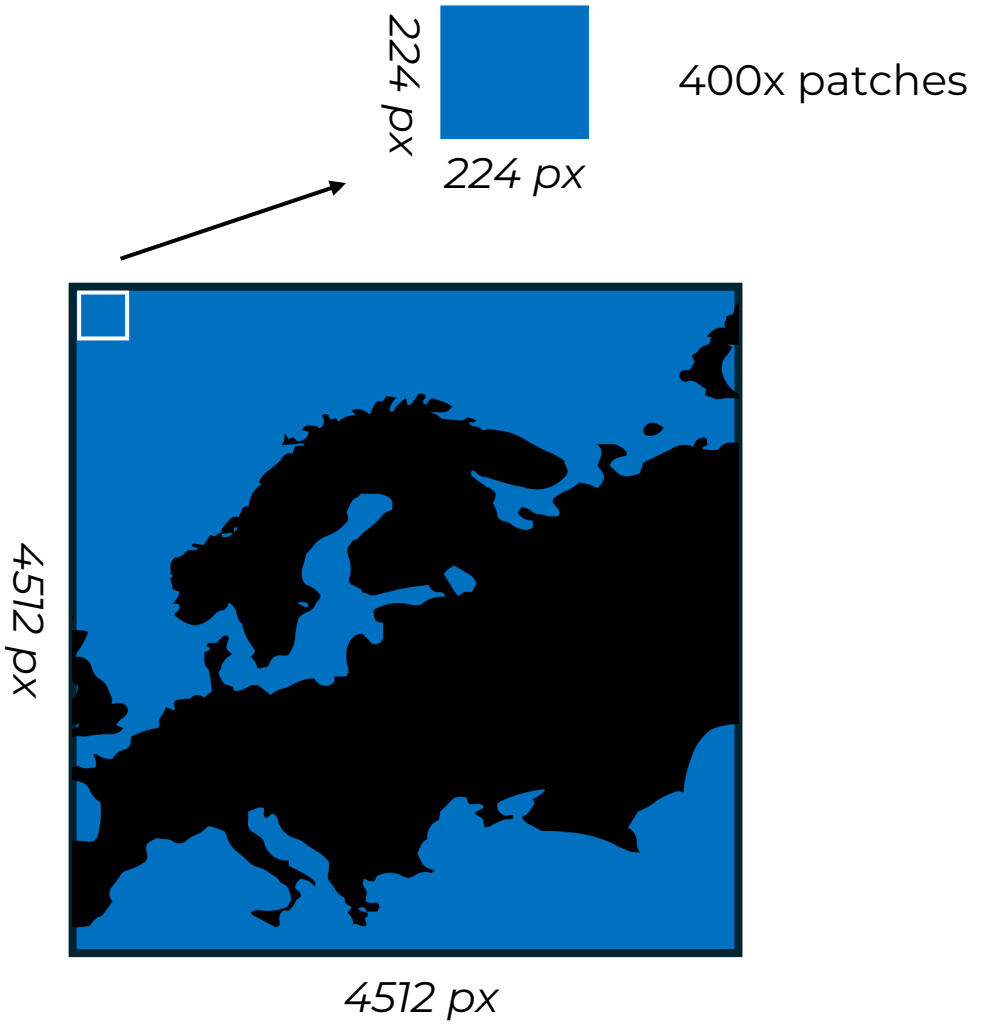
| | ARM Cortex-M7 | Jetson Nano | Toradex Verdin iMX8MP | Coral Edge TPU | Neural Compute Stick 2 |
|---|---|---|---|---|---|
| **Framework** | TensorFlow Lite for Microcontrollers | TensorRT (TensorFlow also tested) | TensorFlow Lite w/ OpenVX backend | Libcoral (TensorFlow Lite w/ Edge TPU runtime) | OpenVino |
| **Quantization** | 8 bit (to fit in memory) | 16 bit | 8 bit (only supports 8 bit ints) | 8 bit (only supports 8 bit ints) | 16 bit (only supports 16 bit floats) |
| **Batching** | Not enough memory to batch | Batch size per 16 (best performing) | Does not support batching | Does not support batching | Number of concurrent inference requests = 4 |

# Model

## Pretrained MobileNetV1

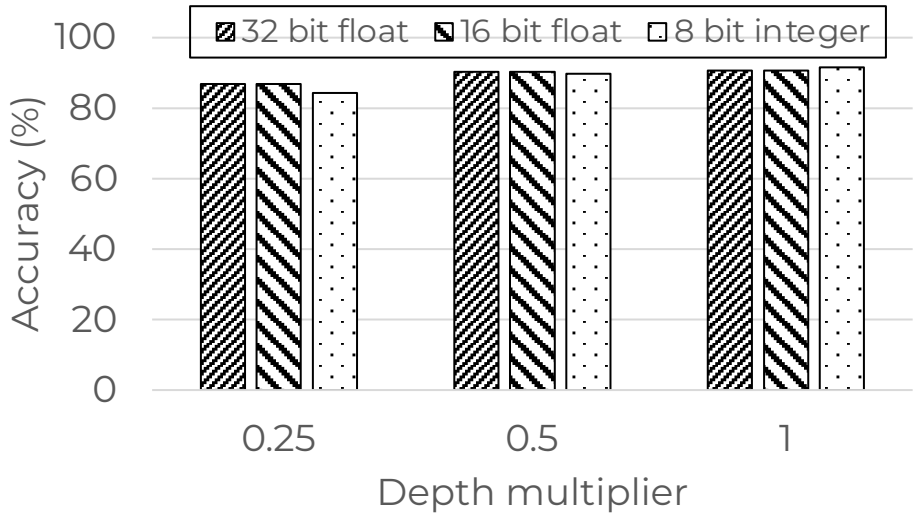



# Data



# Model Considerations

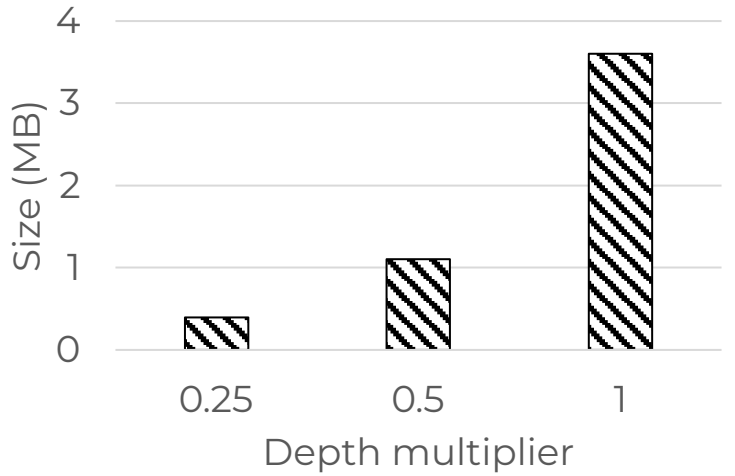
## Depth Multiplier & Quantization




Quantization does not have negative effect on larger models

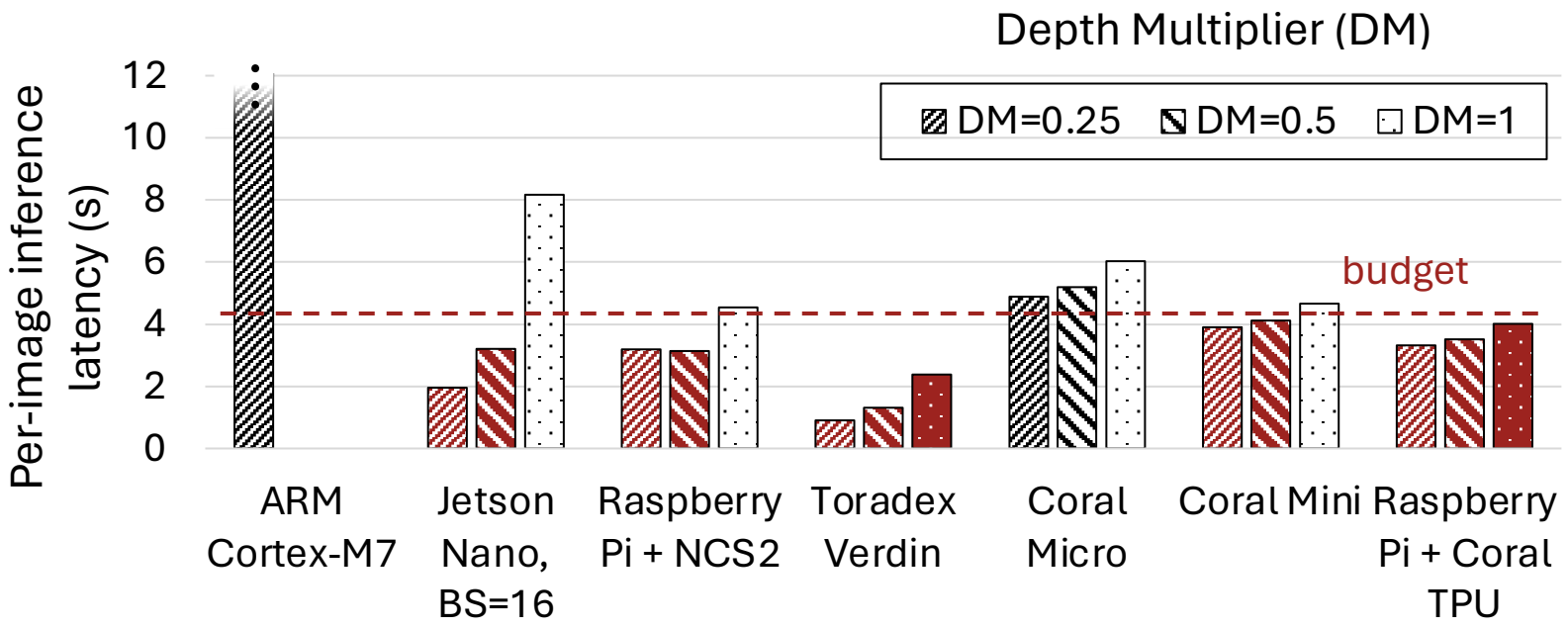
Quantization-aware training can reduce the effect even further

## Storage

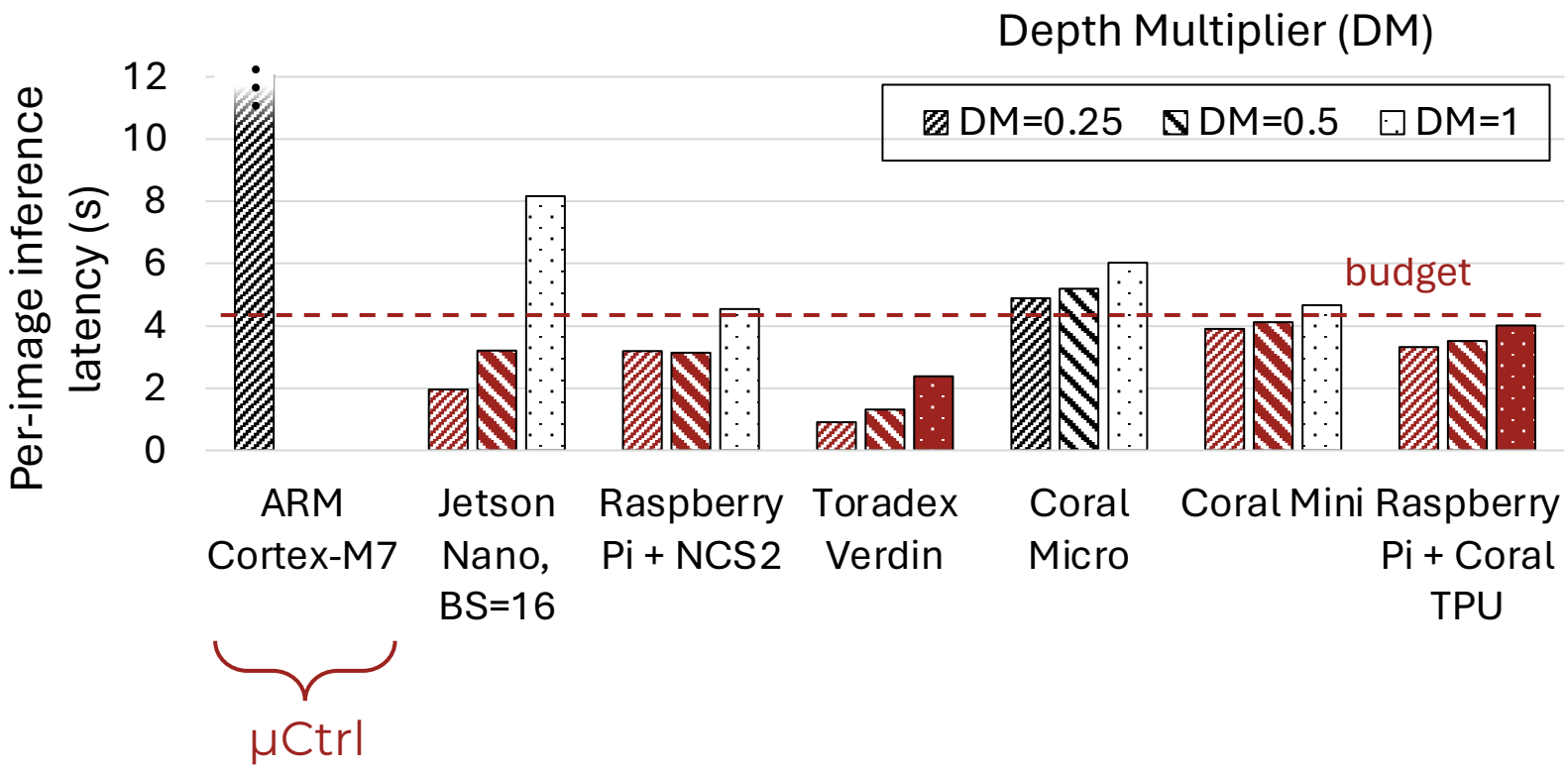



Size of the models is important for updates

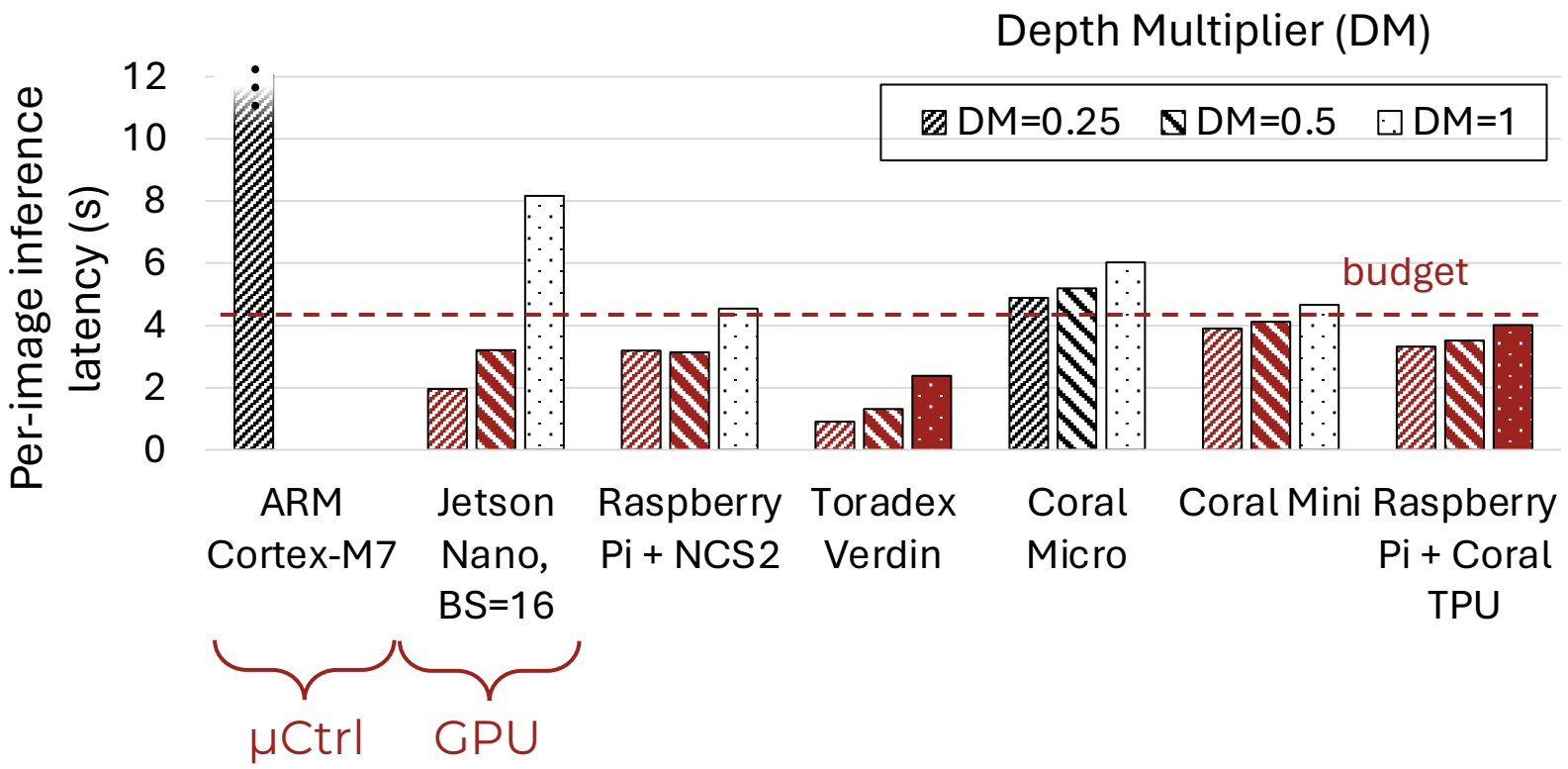
# Results – Latency



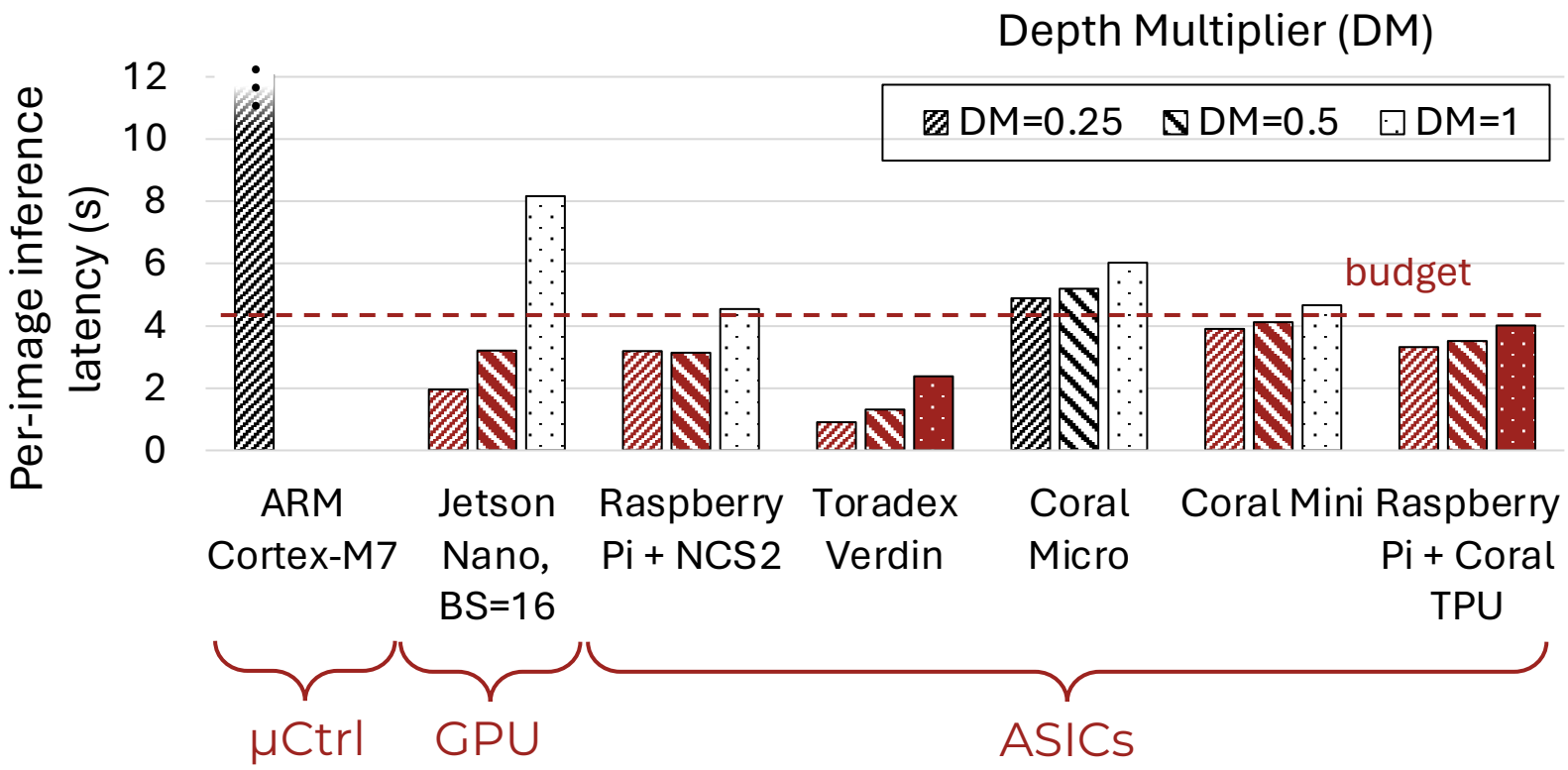
# Results – Latency



# Results – Latency

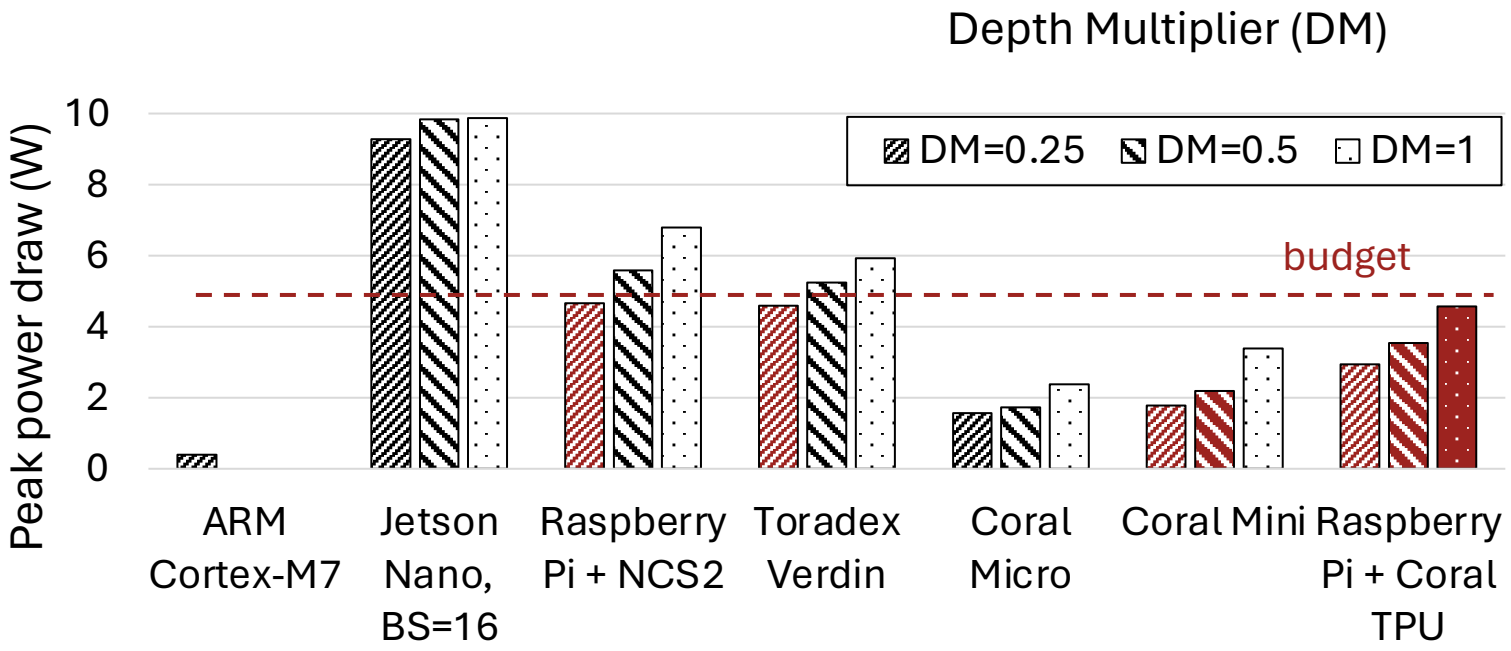


# Results – Latency

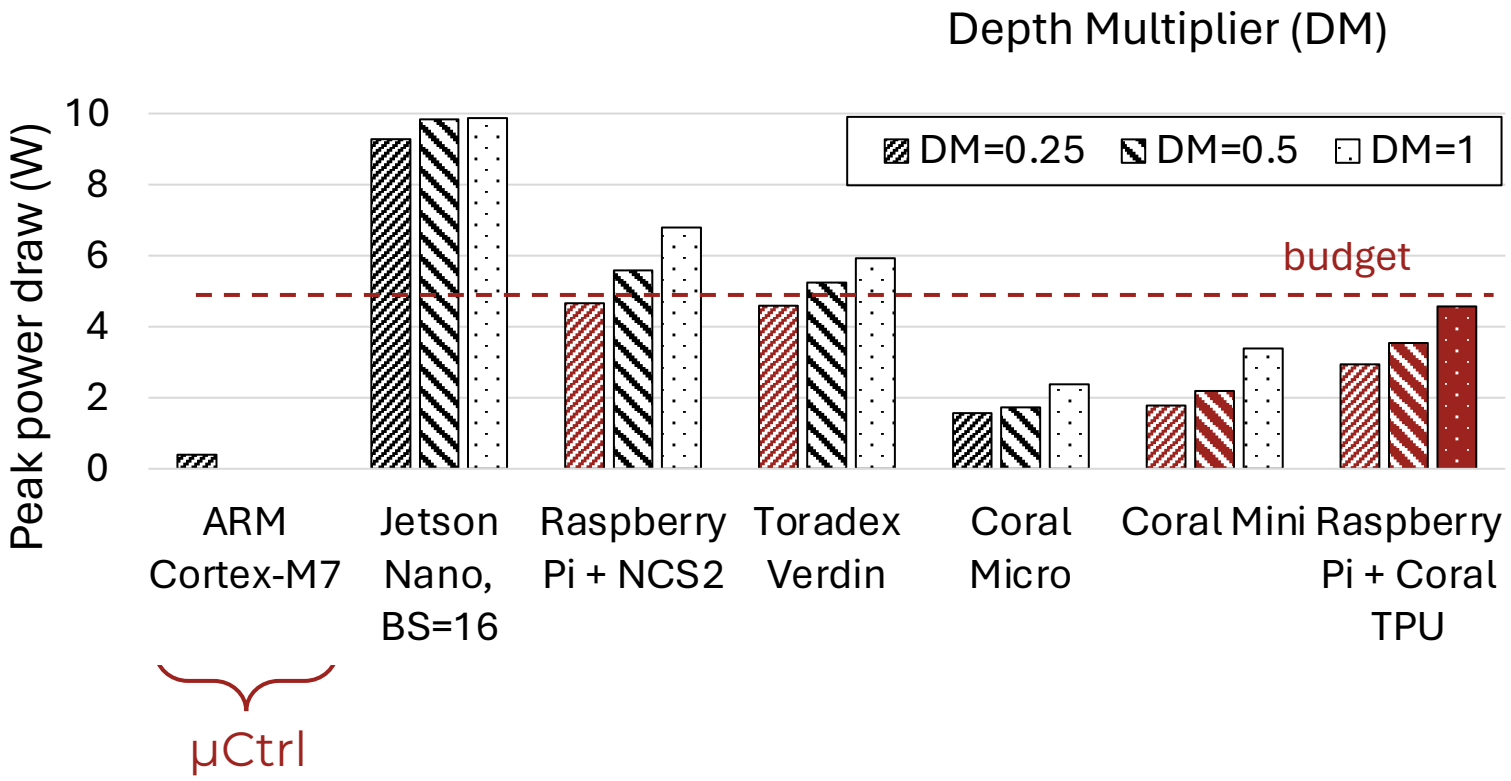



High degree of parallelism necessary

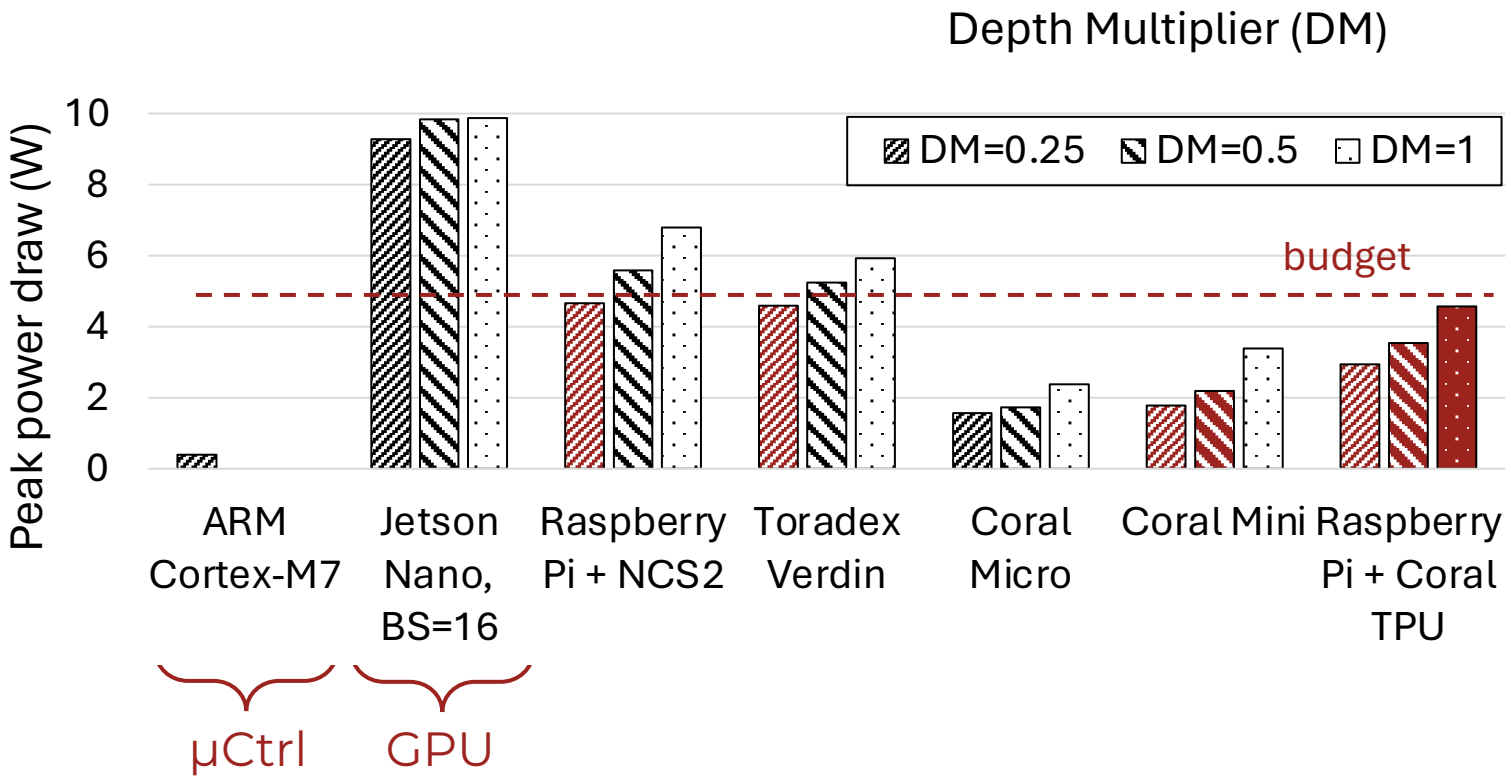
# Results – Peak Power Draw



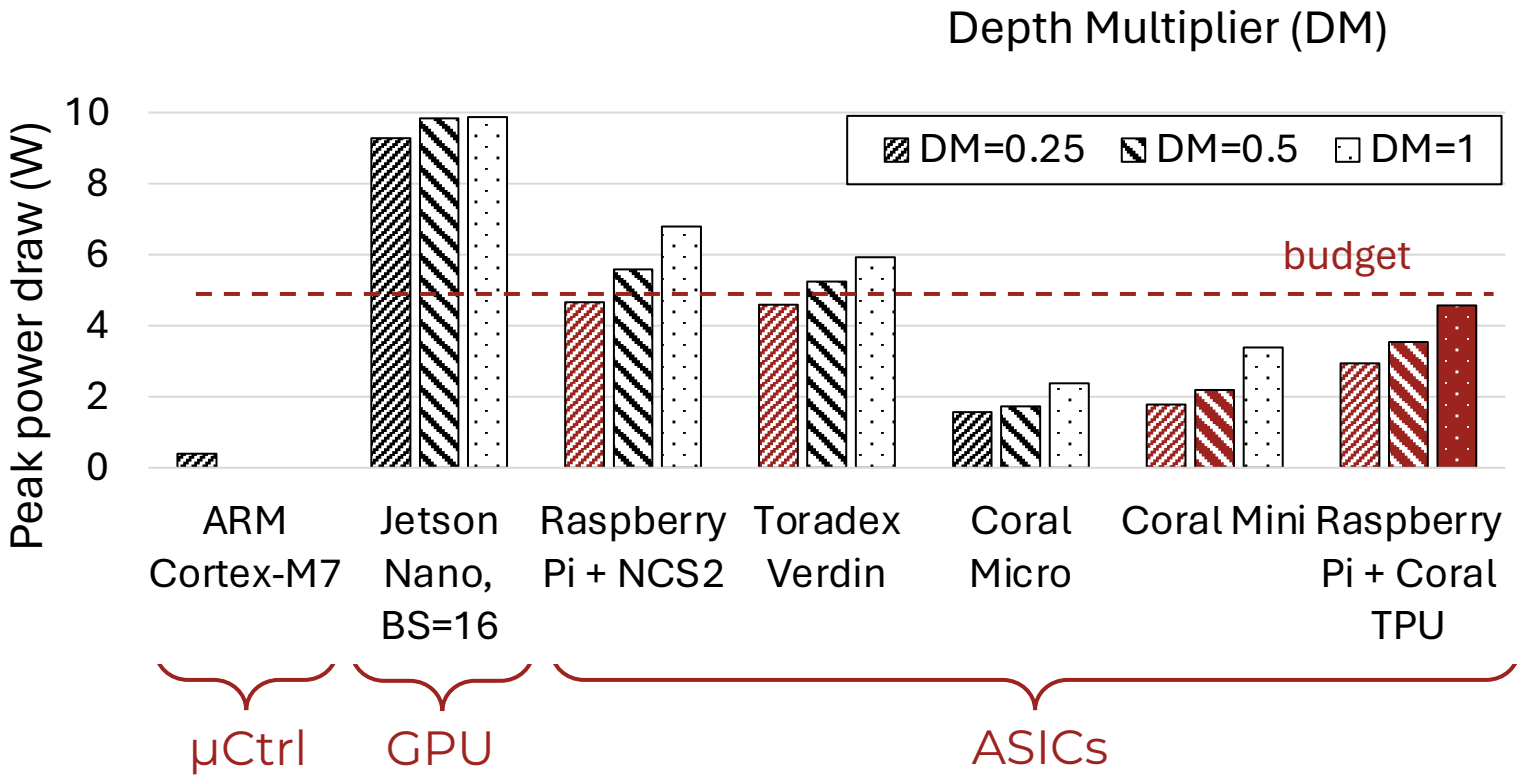
# Results – Peak Power Draw



# Results – Peak Power Draw

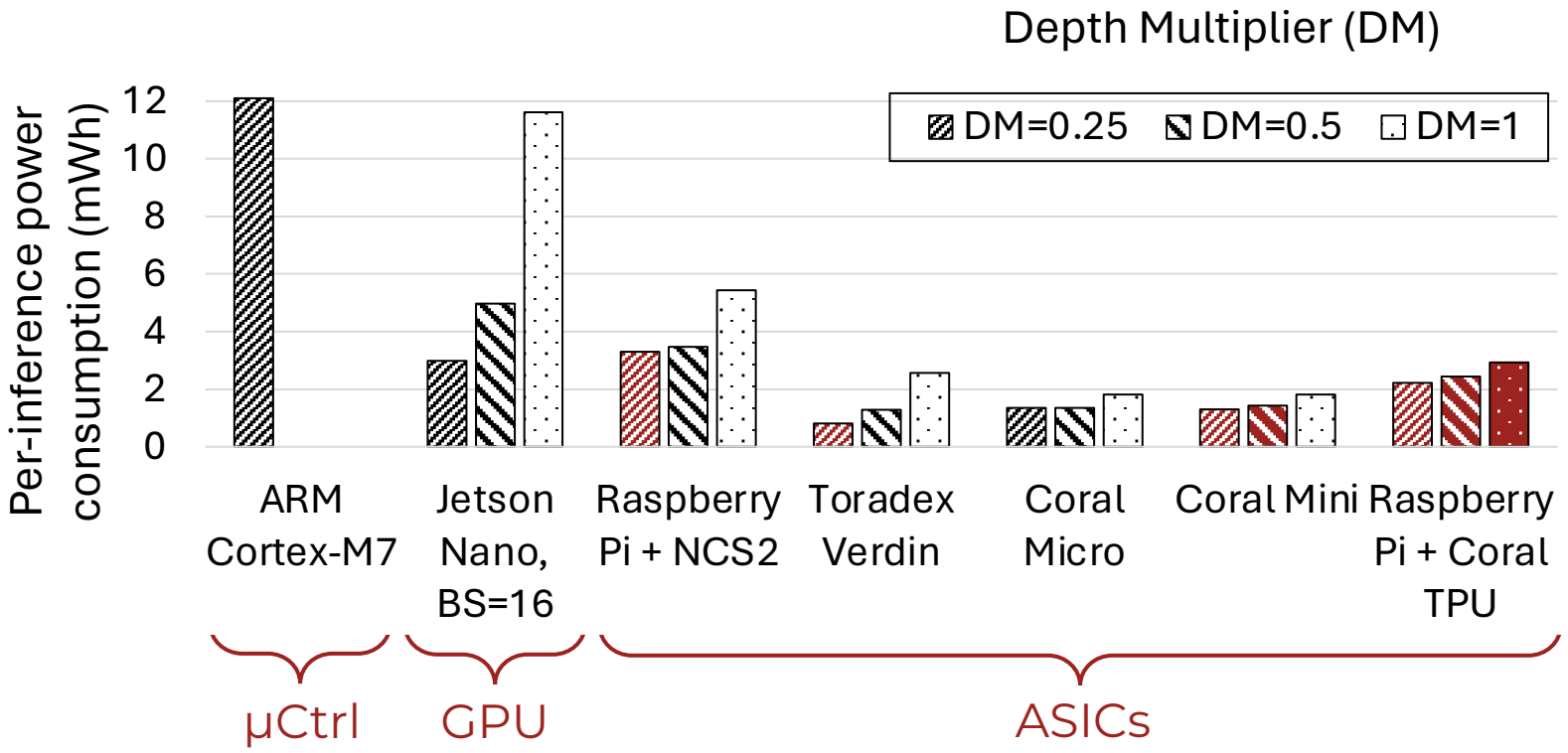


# Results – Peak Power Draw



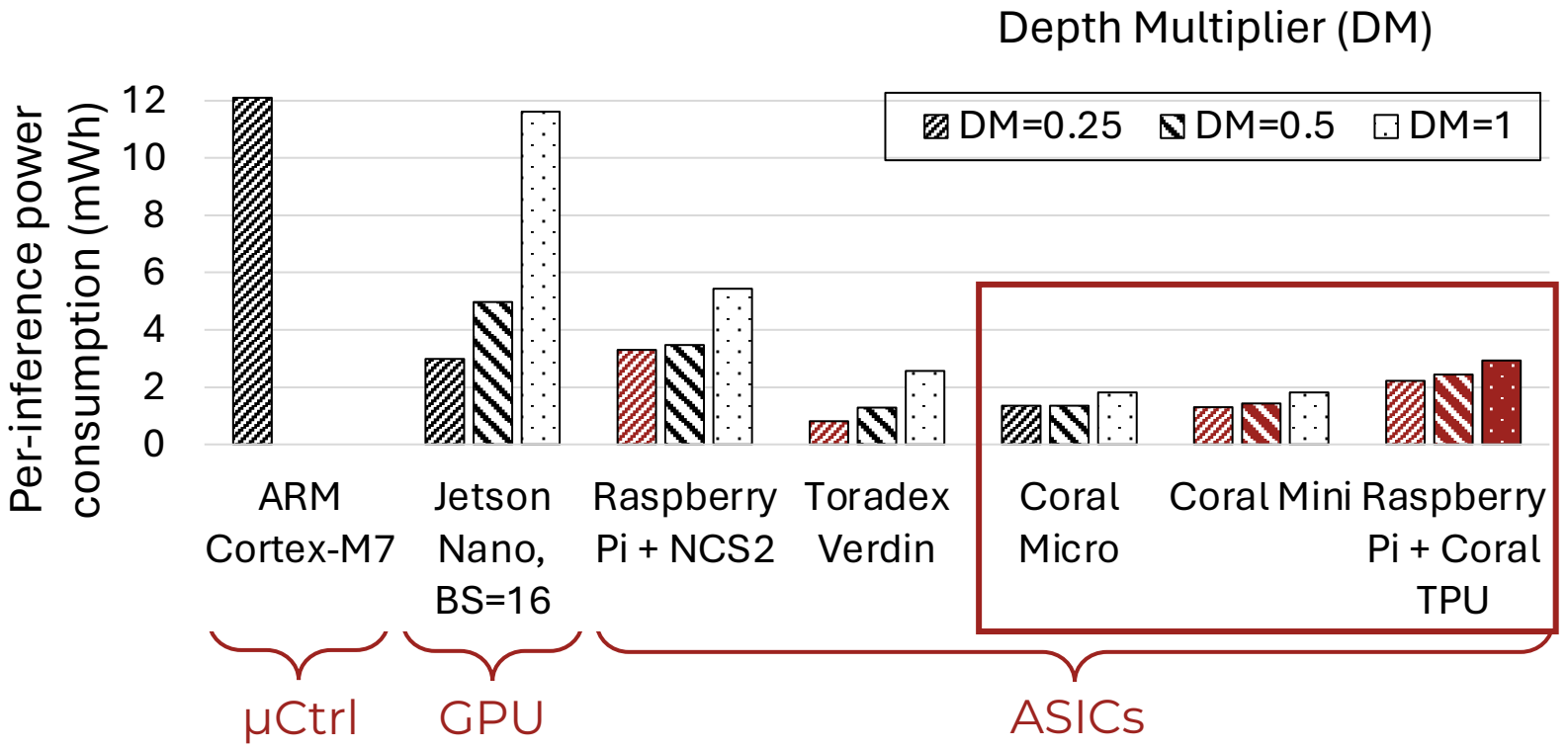

Only the highest specialization can also fulfill power budget

# Results – Power Consumption




The higher the degree of specialization, the higher the efficiency

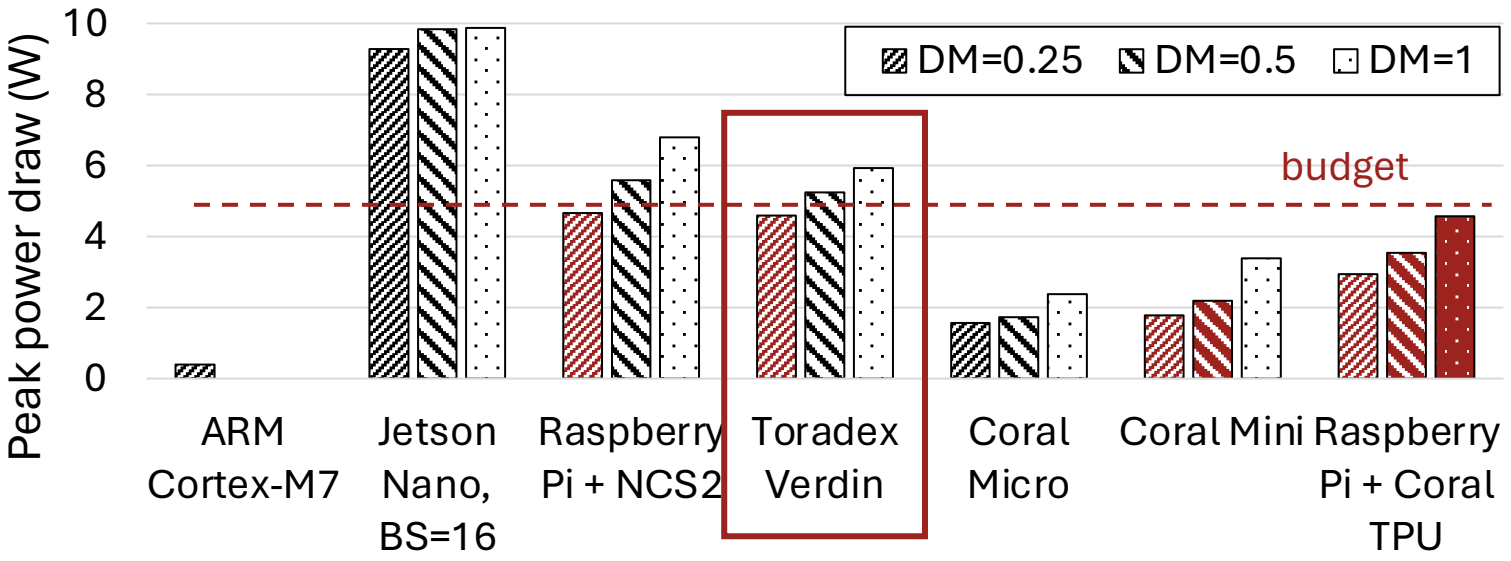
# Results – Power Consumption




The higher the degree of specialization, the higher the efficiency

***Coral Dev Board Mini launched onboard DISCO1 in April 2023***

# DISCO 2




Power measured on development boards

→ Toradex Verdin could fit within the power budget

→ Toradex Verdin achieved best latency and power consumption

# DISCO 2

- Modular, robust, and parametrizable image processing pipeline
- Power-aware scheduling of observations
- Denoising / image stacking
- Data drift detection

# Summary

Data processing at the edge (e.g., on satellites) leads to better network utilization

→ But have to overcome constraints

Highest degree of specialization necessary

→ Provides low latency and power footprint

Coral Dev Board Mini fulfills most requirements

→ Launched on DISCO 1

→ DISCO 2 under development

## Thank you!

GitHub repository

Extended version

Robert Bayer
roba@itu.dk

DEEM
@ SIGMOD

9th June 2024

# Cost Breakdown

- Total (for us): ~$100,000
  - Launch: ~$40,000
  - Parts + engineering (external only): ~$40,000
  - Misc.: ~$20,000

- Fully-commercial: ~$300,000
  - 3U CubeSat: $260,000
  - Launch: ~$40,000