Adaptive Resource Management for CPU and GPU Applications on NVIDIA Jetson TX2 Board

Professor Nikil Dutt

Mentors

Kasra Moazzemi

Bryan Donyanavard

Mentees

Yongtae Kim

Kihoon Park

Byeongkeon Lee

UCIrvine

Index

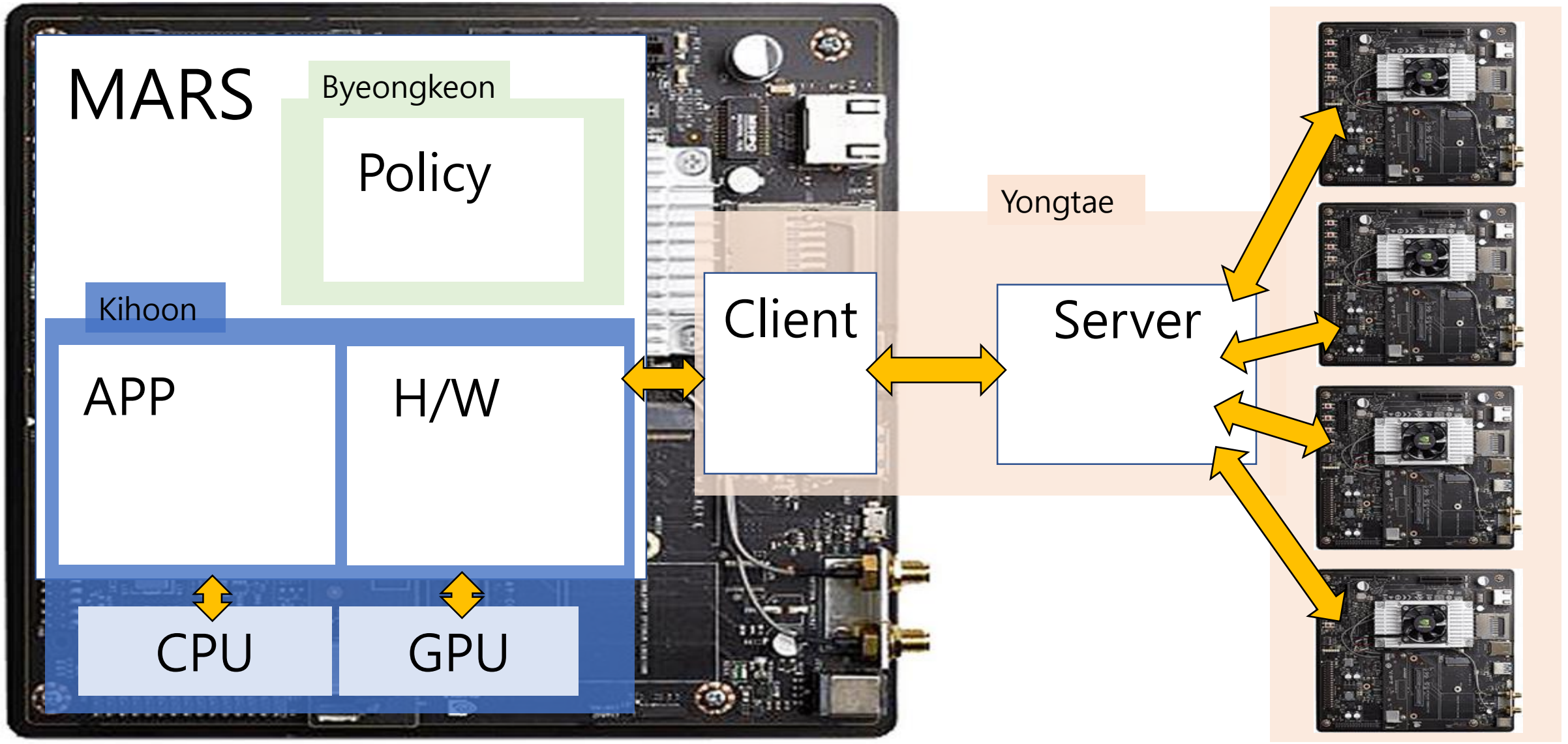
Project Overview.

Project 1. MARS framework development for multi-layer management

Project 2. Adaptive management mechanisms for heterogeneous platforms

Project 3. Heartbeat annotation and measurement of modern benchmarks on NVIDIA Jetson TX2 platform

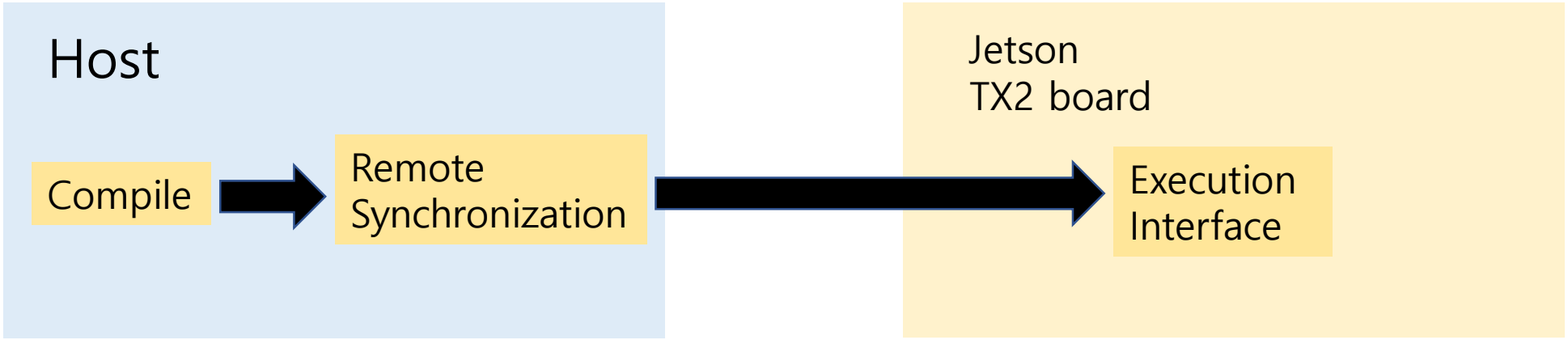
Project Overview



MARS framework development for multi-layer management

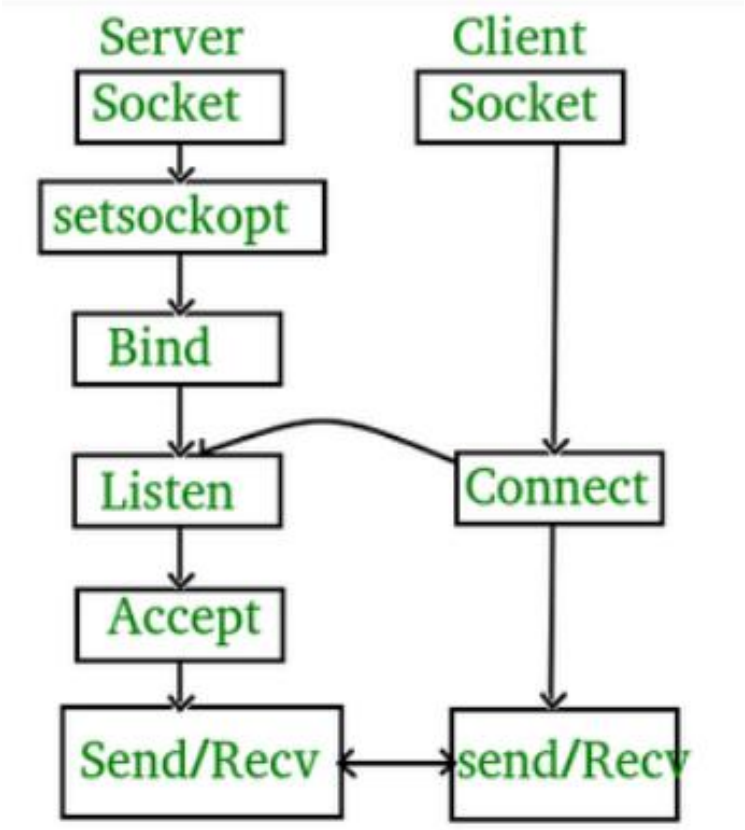
1. Study and Summarize of MARS Paper

2. Cross compile from the host computer to Jetson TX2 board



MARS framework development for multi-layer management

3. Using TCP to implement interactive programs
server(host computer) and local(Jetson TX2 board)

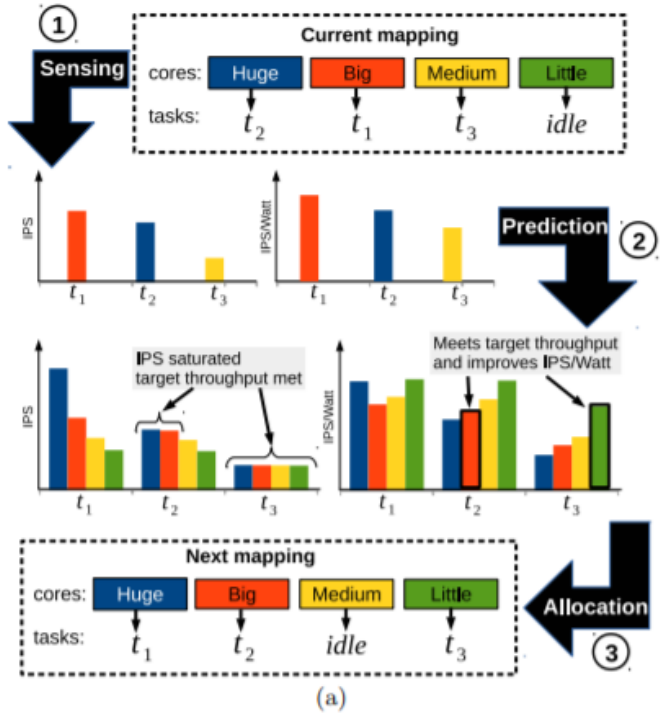


Adaptive management mechanisms for heterogeneous platforms

- Reference

SPRATA : Heterogeneous Task allocation

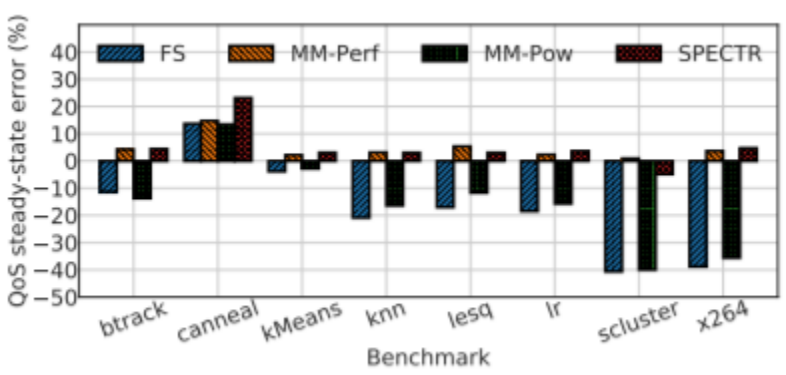
- Runtime sensing : Run all tasks (1 epohcs)
- Prediction : Predict how much power will be consumed, how much time will be executed
- Task allocation : difference between IO-intensive job and computation integrated job.

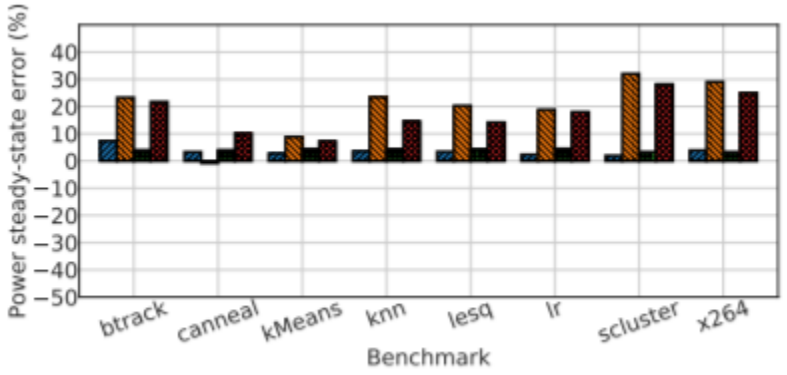


Adaptive management mechanisms for heterogeneous platforms

- Reference

SPECTR : Evaluation method

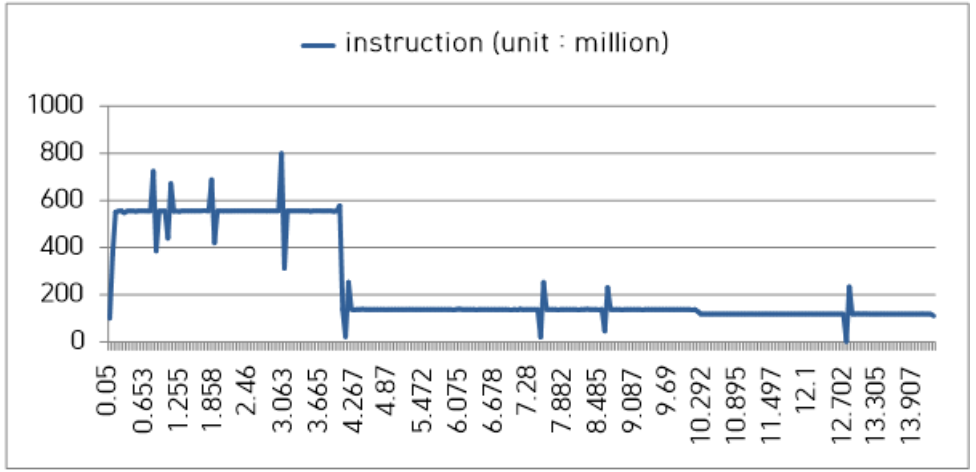


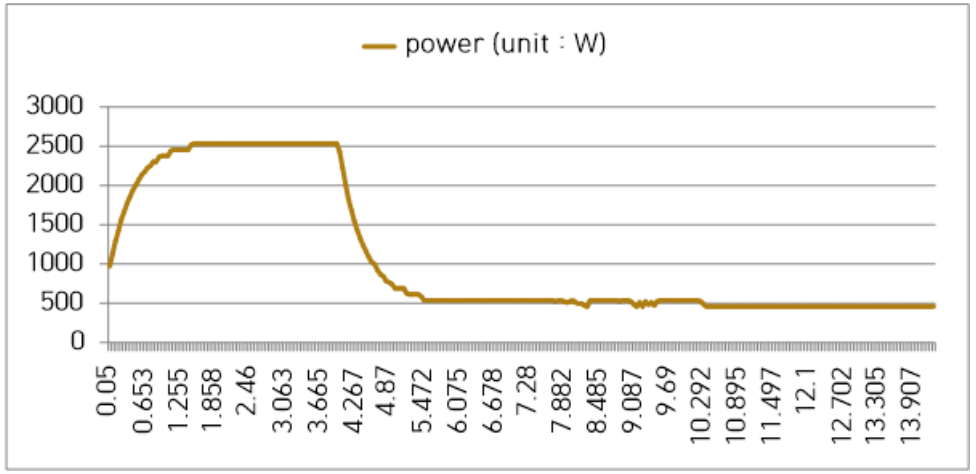


Adaptive management mechanisms for heterogeneous platforms

- Current Interface test

sample_id	timestamp	freq_domair	freq_domair	freq_domair	total_instr	freq_domair	total_cpu_ti	total_power
0	0,05	960	960	2035	99584500	2035	0,087	972,3333
1	0,1	960	960	2035	3,722E+8	2035	0,11	1132,625
2	0,15	960	960	2035	5,513E+8	2035	0,127	1294
3	0,201	960	960	2035	5,546E+8	2035	0,133	1423,778
4	0,251	960	960	2035	5,565E+8	2035	0,127	1569,222
5	0,301	960	960	2035	5,479E+8	2035	0,127	1669
6	0,351	960	960	2035	5,55E+8	2035	0,129	1776
7	0,402	960	960	2035	5,555E+8	2035	0,127	1859,111
8	0,452	960	960	2035	5,553E+8	2035	0,131	1944,333





Adaptive management mechanisms for heterogeneous platforms

- To do

- Reinforcement Learning paper

F. M. M. u. Islam and M. Lin, "Hybrid DVFS Scheduling for Real-Time Systems Based on Reinforcement Learning"

- Analyze MARS more and code modification

to improve task allocation with reinforcement Learning

Mars-benchmarks

- What is benchmark?

  Program for measure the relative **performance** of an object to improve **energy efficiency**

- PARSEC etc. in mars-benckmarks are only for CPU now

- Extending this to include GPU workloads

Jetson TX2 Board

Mars-benchmarks

PARSEC
- bodytrack
- x264
- …

…

CPU

+

Rodinia 3.1 (benchmark)
- cuda
  - gaussian
  - …

GPU

video processing
neural networks.

what Expect?

1) Determine running frequency of GPU cores

2) Measure memory utilization

3) Evaluate the conflicts in resources

Thank you for Listening

Q&A