ADOPTING MANYCORE AS AUTOWARE ACCELERATOR

Stephane Strahm
Sr Product Marketing Mgr, Kalray
ACCELERATION COMPUTE
ACCELERATION COMPUTE

MAIN FUNCTIONS

• Awareness of environment
• Decision making
• Compute capability
• Algorithms acceleration
• Sensor raw data analysis
• Image analysis
• Path planning
ACCELERATION COMPUTE

MAIN FUNCTIONS

• Awareness of environment
• Decision making
• Compute capability
• Algorithms acceleration
• Sensor raw data analysis
• Image analysis
• Path planning

ESTIMATED PERFORMANCE

• L2: 10-20 TOPS
• L3: 20-100 TOPS
• L4-5: 100-300+ TOPS
MAIN FUNCTIONS

- Awareness of environment
- Decision making
- Compute capability
- Algorithms acceleration
- Sensor raw data analysis
- Image analysis
- Path planning

ESTIMATED PERFORMANCE

- L2: 10-20 TOPS
- L3: 20-100 TOPS
- L4-5: 100-300+ TOPS

KEY PRODUCTION CRITERIA

- Performance per Watts
- Scalability
ACCELERATION COMPUTE

MAIN FUNCTIONS

• Awareness of environment
• Decision making
• Compute capability
• Algorithms acceleration
• Sensor raw data analysis
• Image analysis
• Path planning

ESTIMATED PERFORMANCE

• L2: 10-20 TOPS
• L3: 20-100 TOPS
• L4-5: 100-300+ TOPS

MANDATORY ELEMENTS

Safety
Security

KEY PRODUCTION CRITERIA

• Performance per Watts
• Scalability
ACCELERATOR - SYSTEM INTEGRATION

SENSORS
- CAMERA
- RADAR
- LIDAR
- GPS
- OTHERS

PERCEPTION
- LOCALIZATION
- DETECTION
- TRACKING
- CLASSIFICATION
- SEGMENTATION

PLANNING
- ROUTE PLANNING
- PREDECTION
- BEHAVIOR PLANNING
- TRAJECTORY PLANNING

CONTROL
- ENGINE
- STEER
- BRAKE
- OTHERS

HOST PROCESSOR(S)

SENSOR I/F

CONTROL I/F

PCle

PCle
AUTOWARE HIGH-LEVEL MODULES

Source: https://github.com/CPFL/Autoware/wiki/Overview
- Perception module: Heaviest algorithms
  - Image detector – based on CNN (Yolo, SSD, R-CNN)
  - Fusion detector – based on DNN (MV3D)
  - LiDAR localizer – based on NDT Matching
  - LiDAR detector – based on DNN (VoxelNet, LMNet)

- Decision, Planning: compute algorithms can be offloaded, integration design
AUTOWARE PERCEPTION MODULE

AVAILABLE FUNCTIONS ON MPPA®

- Image Detector
  - CNN based object detection (car/person/…)
  - Autoware uses Darknet(YOLO networks)/SSD/RCNN project
  - SSD and Yolo V3 CNN have been tested within Autoware, using Kalray inference engine solution on MPPA®
  - Other CNN can also be used

- LiDAR Localization
  - Based on “NDT matching” algorithm from PCL library
  - Compute car coordinates by matching its LIDAR data to a precomputed map
  - Fully implemented on MPPA® (both initialization and runtime functions)
IMAGE DETECTOR

Camera

Vision YOLOv3 Detect

Darknet API

Darknet YOLOv3

GPU

detectedCar

detectedPerson

to Fusion/Planning/Decision

/imageRaw
IMAGE DETECTOR

Camera

/ImageRaw

Vision YOLOv3 Detect

/(detectedCar)

/(detectedPerson)

Forward Service
to Fusion/Planning/Decision

MPPA® Service (YOLOv3)

MPPA®

MPPA API
LIDAR LOCALIZATION

LIDAR

IMU

GPS

/IMUraw

/PointsRaw

/gnssPose

NDT Matching

PCL impl

GPU

to Fusion/Planning/Decision

/estimated...
LIDAR LOCALIZATION

- LIDAR
- IMU
- GPS

/IMUraw
/PointsRaw
/gnssPose

NDT Matching

MPPA® Service (NDTM)

NDTM Forward Service

MPPA®

to Fusion/Planning/Decision

MPPA API

© 2019 Kalray SA All rights reserved. Confidential information
OK, IT WORKS!

PRESENTATION SHOULD INCLUDE VIDEO

REMOVED HERE FOR FORWARDING BY EMAIL
MANYCORE MPPA®

Porting on Manycore MPPA®

- Smooth integration, Same implementation… It works
- Keeping Autoware software infrastructure… Facilitates maintenance

Adopting Manycore MPPA®

- Scalability: Over MPPA® clusters (5 clusters w/16 cores each, multiple dies up to 160 cores)
- Isolation: No interference between acceleration modules (HW isolation by design)
- Consolidation: Simultaneous execution of multiple acceleration modules on single chip
- Determinism: Execution time control (VLIW patented architecture) Distributed memory, NOC
- Performance: 3rd Generation (MPPA®3) in 2019 at 25 TOPS
- Consumption: Embedded TOPS/Watt level
- Development: Use Open Standards (C/C++/GCC/GDB/LLVM/OpenCL/Eclipse)
NEXT STEPS

Moving market to benefit from Autoware Eco-System
• From Experimentation to Reference Implementation to Production

Autoware members have all bricks for Reference Implementation Step
• Experimentation Step is on-going
• Let’s move to Reference Implementation Step
• Define it, plan it, make it
• All Autoware members have a brick
NEXT STEPS

Moving market to benefit from Autoware Eco-System

• From Experimentation to Reference Implementation to Production

Autoware members have all bricks for Reference Implementation Step

• Experimentation Step is on-going
• Let’s move to Reference Implementation Step
• Define it, plan it, make it
• All Autoware members have a brick

Use Cases: Autoware.Auto

Requirements: partial

Architecture: in progress

Implementation: in progress

Validation: Simulator

Tests: status

Release: To be defined
Moving market to benefit from Autoware Eco-System

- From Experimentation to Reference Implementation to Production

Autoware members have all bricks for Reference Implementation Step

- Experimentation Step is on-going
- Let’s move to Reference Implementation Step
- Define it, plan it, make it
- All Autoware members have a brick

Use Cases: Autoware.Auto
Requirements: partial
Architecture: in progress
Implementation: in progress
Validation: Simulator
Tests: status
Release: To be defined

Software Framework
Autoware.Auto and Autoware.IO
Reference Hardware
Under definition
Reference Accelerator Boards

- PCIe boards with MPPA®2 Generation, MPPA®3 upcoming
- Accelerator daughter board modules

Reference Programming Software

- CNN Inference Optimizer
- CNN Inference Runtime Engine (additional algo support and growing)
- Computer Vision OpenCV
- OpenCL 1.2

Hardware Integration

- Intel and ARM hosts

Software Integration

- Additional Use Cases support (Companion, StandAlone)-to simplify System integration
- Linux and Autoware Member RTOS supported
THANK YOU

IV2019 – © KALRAY

www.kalrayinc.com