KALRAY

Kalray MPPA®

Massively **P**arallel **P**rocessor **A**rray

The MPPA®-256 Bostan Manycore Processor for Real-Time and Safety-Critical Systems

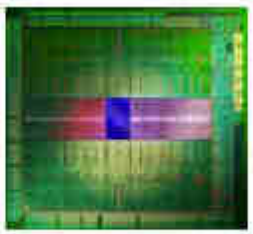
Benoît Dupont de Dinechin, CTO

Landscape of Computing Technologies

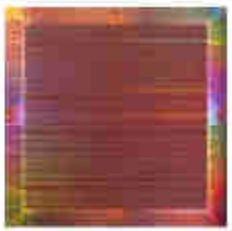
- Field-Programmable Gate Arrays (FPGA)
  - Most effective on bit-level computations
  - Require HDL programming
  - **Suitable for safety-critical computing**
- Graphics Processing Units (GPU)
  - Most effective on regular computations
  - Require CUDA or OpenCL programming
  - **Unsuitable for safety-critical computing**
- Digital Signal Processors (DSP)
  - Most effective on fixed-point arithmetic
  - Require low-level programming
  - **Suitable for safety-critical computing**
- Intel Many Integrated Core (MIC)
  - Require multicore programming + exploitation of SIMD instructions (AVX)
  - **Unsuitable for safety-critical computing**

CPUs

DSPs



GPUs



FPGAs

Courtesy Altera

Safety-Critical Computing

- Time-critical requirements
  - Time constraints associated with information manipulation
  - Execution time determinism, predictability, composability
  - Certification of worst-case response time by static analysis
- Dependability requirements
  - Detect, isolate or mitigate errors that occur in digital circuits
  - Transient electrical problems, soft errors, physical damage
- Execution timing issues
  - Intra-core non-deterministic timing: speculation, branch prediction
  - Inter-core competition for shared resources: caches, busses, devices
  - Ordering of task executions in critical sections
- Application domains considered
  - Aerospace & defense; autonomous vehicles;
  - Financial trading; large-scale physical instrumentation;
  - industrial robotics; manufacturing equipment;

MPPA® MANYCORE Processor Roadmap

2014	2015	2016	2017

MPPA®-1024

Volume

Samples Volume

Samples Volume

MPPA®-256

MPPA®-64

TSMC 28nm (HP)		28nm/16nm
Andey	**Bostan**	**Coolidge**
Kalray 1st generation	**Kalray 2nd generation**	**Kalray 3rd generation**
211 GFLOPS SP	634 GFLOPS SP	1056 GFLOPS SP
70 GFLOPS DP	316 GFLOPS DP	527 GFLOPS DP
32-bit addressing	64-bit addressing	64-bit addressing
400Mhz	600Mhz	1000Mhz

Production

Development

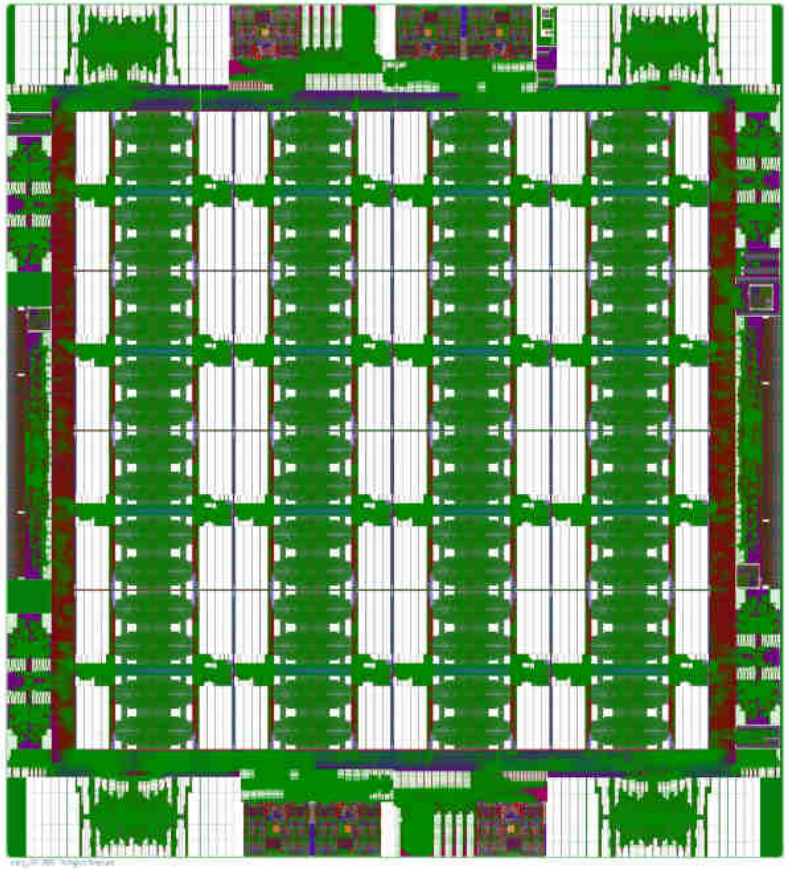
Under Specification

MPPA® MANYCORE Architecture Highlights

- DSP type of acceleration
  - Energy efficiency
  - Timing predictability
  - Software programmability
- CPU ease of programming
  - C/C++ GNU programming environment
  - 32-bit or 64-bit addresses, little-endian
  - Rich operating system environment
- Integrated many-core processor
  - 32 management cores on chip
  - 256 application cores on chip
  - High-performance low-latency I/O
- Scalable massively parallel computing
  - MPPA® processors tiled together through NoC extensions

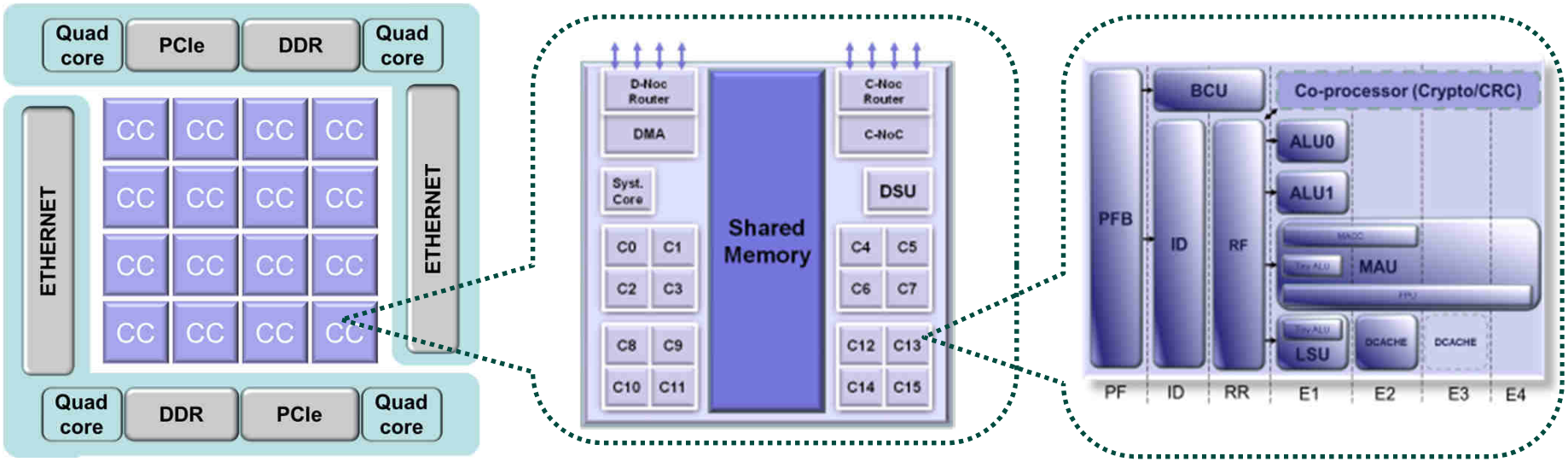
MPPA®-256 Bostan Processor

256 + 32 VLIW cores / 18 address spaces / 2D Torus dual NoC



- Physical characteristics
  - TSMC CMOS 28HP
  - 100µW/MHz per core + L1 caches
  - 2W to 3W leakage
- Processor interfaces
  - 2x DDR3 Memory interfaces
  - 2x PCIe Gen3 8-lane interface
  - 8x 1G/10G or 2x 40G Ethernet interfaces
  - SPI/I2C/UART interfaces
  - Universal Static Memory Controller (NAND/NOR/SRAM)
  - GPIOs with Direct NoC Access
  - NoC extension through Interlaken interface (NoCX)

MPPA®-256 Bostan Processor Architecture



Manycore Processor

- 16 compute clusters
- 2 I/O clusters each with quad-core CPUs, DDR3, 4 Ethernet 10G and 8 PCIe Gen3
- Data and control networks-on-chip
- Distributed memory architecture
- 634 GFLOPS SP for 25W @ 600Mhz

Compute Cluster

- 16 user cores + 1 system core
- NoC Tx and Rx interfaces
- Debug & Support Unit (DSU)
- 2 MB multi-banked shared memory
- 77GB/s Shared Memory BW
- 16 cores SMP System

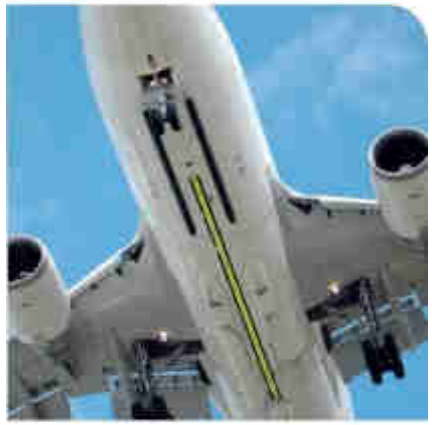
VLIW Core

- 32-bit or 64-bit addresses
- 5-issue VLIW architecture
- MMU + I&D cache (8KB+8KB)
- 32-bit/64-bit IEEE 754-2008 FMA FPU
- Tightly coupled crypto co-processor
- 2.4 GFLOPS SP per core @600Mhz

MPPA® Processor Co-Design for Avionics

- U. Saarland / AbsInt GMBH recommendations on VLIW core and cache micro-architecture design
  - AbsInt provides the aiT static timing analysis tool used to certify the flight control system of Airbus A380, Airbus A350 and Airbus A400M
  - AbsInt aiT tool also targets the Kalray VLIW cores
- Architecture with a focus on timing predictability
  - Core level: micro-architecture
    - ✓ Fully timing compositional core
    - ✓ LRU caches and write buffer
    - ✓ Cache bypass memory accesses
  - Cluster level: multi-banked shared memory
    - ✓ Core-private buses for memory bank access
  - Processor level: NoC with guaranteed services
    - ✓ Minimum bandwidth & maximum latency

CERTAINTY



Certification of Real Time Applications desigNed for mixed criticaliTY

THALES

UNIVERSITE JOSEPH FOURIER

ARTTIC

AbsInt

UPPSALA UNIVERSITET

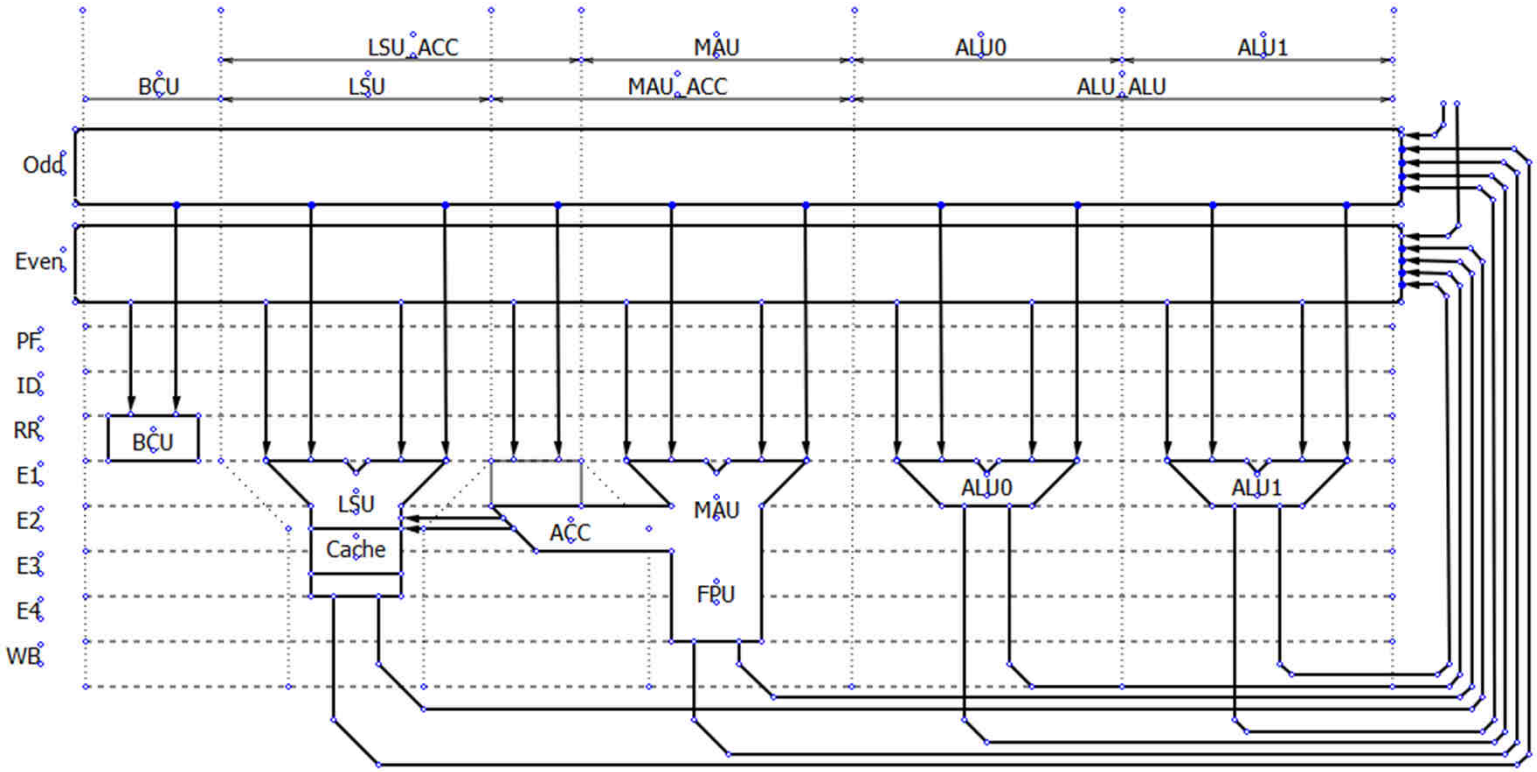
ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

Kalray VLIW Architecture Compared to HP Labs Lx

The Lx architecture begat the STMicroelectronics ST200 VLIW family

- Optimize use of the data memory bandwidth
  - Widening to 64-bit, no alignment restrictions
  - Enable large immediate values in instruction stream
  - All memory accesses may bypass the L1 data cache & write buffer
- Eliminate DLX ISA features and restrictions
  - Instructions with 3 or 4 source operands, 1 or 2 target operands
  - No aliasing between registers and special resources (LR, zero)
  - Memory addressing modes similar to those of PowerPC
  - Effective floating-point support with Fused Multiply Add
- Rework if-conversion support
  - Remove Boolean registers and SELECT instructions
  - Use CMOV and conditional load/store instructions
- Support hardware looping

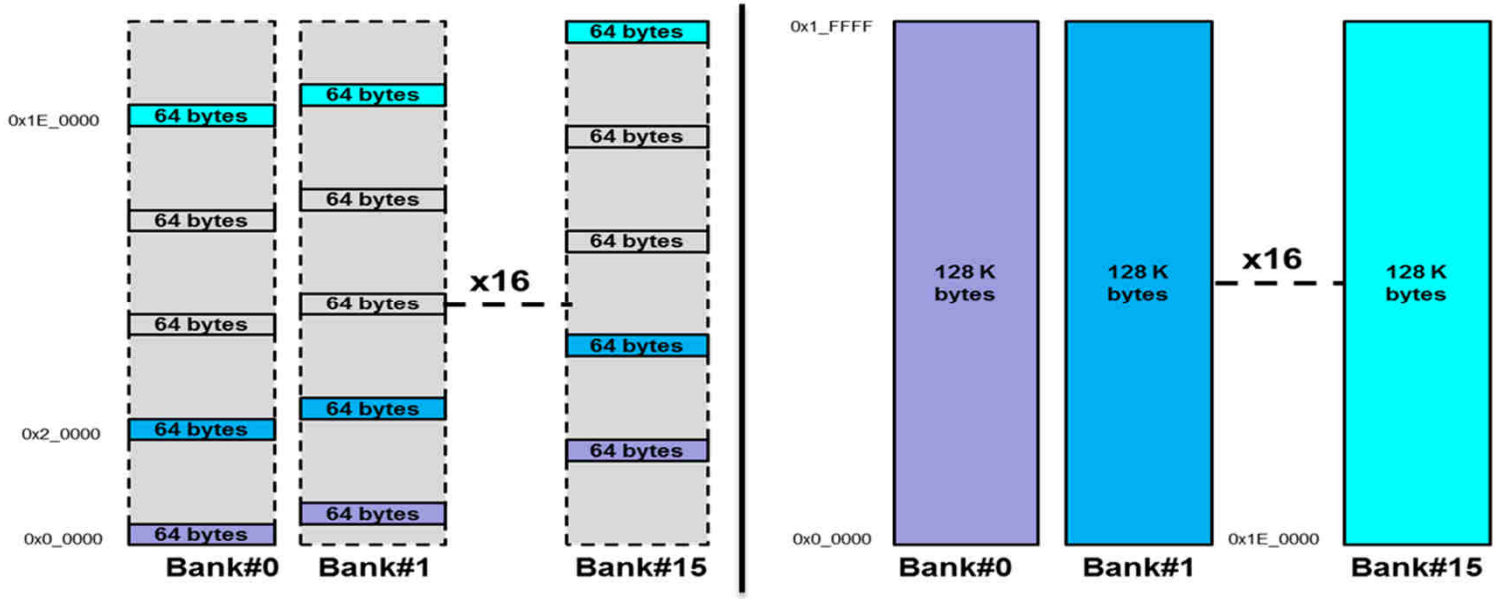
MPPA®-256 Bostan VLIW Core Data Path



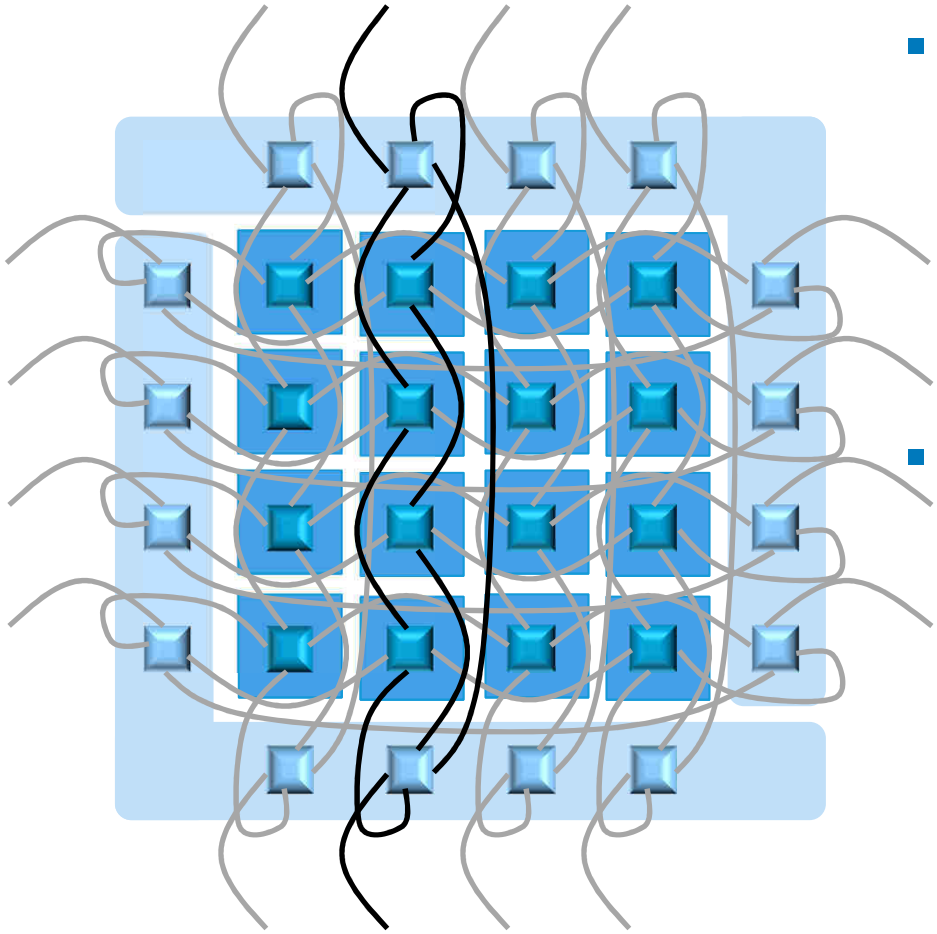
- 5-issue, polycyclic property
- Unified 10 read, 5 write 64x32-bit register file
- 64-bit data in register pairs
- Shared ACC read port
- Unified MAU and FPU
- LSU and MAU also execute single-cycle ALU instructions

MPPA®-256 Bostan Compute Cluster

- 20 bus masters
  - 16 application cores
  - 1 management core
  - NoC Tx and Rx interfaces
  - Debug support unit (DSU)
- 16-banked shared memory
  - 2MB extensible to 4MB
  - No bus interferences between cores
  - RR arbitration between bus masters
  - Interleaved or blocked address map

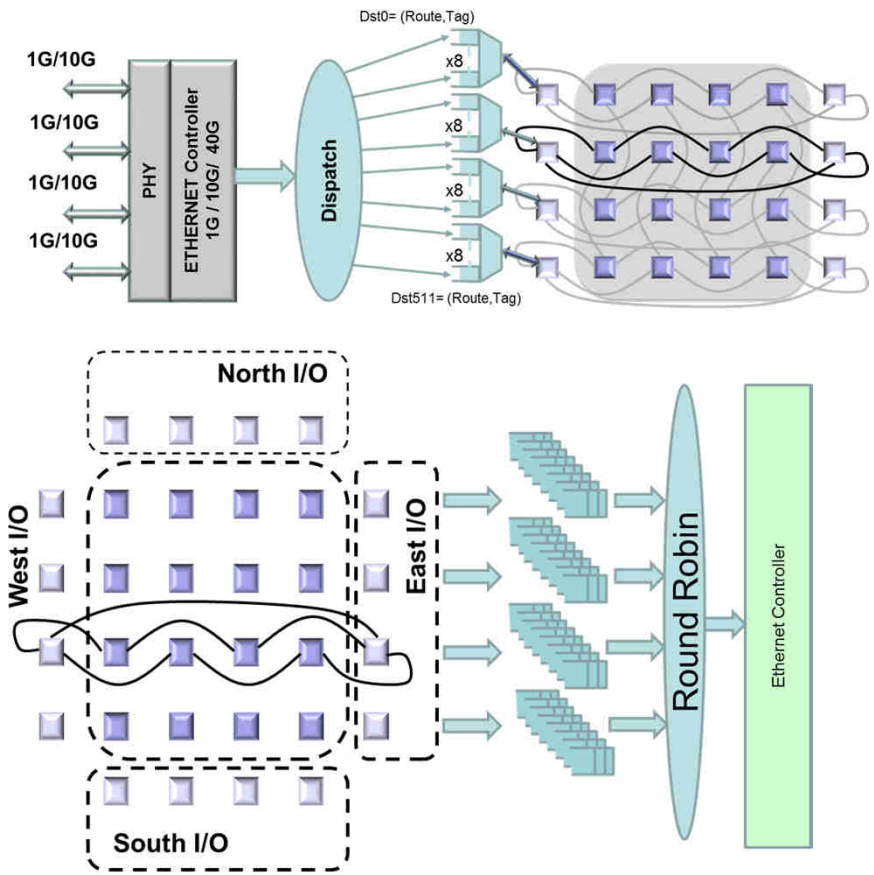


MPPA®-256 Bostan Network-on-Chip (NoC)



- Dual 2D-torus NoC
  - D-NoC: high-bandwidth RDMA
  - C-NoC: low-latency mailboxes
  - 4B/cycle per link direction per NoC
  - Nx10Gb/s NoC extensions for connection to FPGA or other MPPA®
- Predictability
  - Data NoC is configured by selecting routes and injection parameters
  - Injection parameters are the (σ,ρ) or (burst, rate) of Cruz network calculus
  - Guaranteed services rely on same methods as in AFDX Ethernet

MPPA®-256 Bostan Ethernet Support



- Ethernet as the main high-performance / low-latency IO
  - Integration of Ethernet Rx and Tx to the D-NoC architecture
- Per Ethernet Rx port
  - 8 classification tables for hardware dispatch
  - Round-robin or classified cluster & core allocation
- Per Ethernet Tx port
  - 64 independent Tx FIFOs
  - Weighted round-robin between Tx FIFOs
  - Flow control between clusters and Tx FIFOs

MPPA® AccessCore™ SDK

Standard C/C++
Programming Environment

Simulators & Profilers,
Debuggers & System Trace

Operating Systems
& Device Drivers

AccessLib™
Optimized Libraries

C - Low Level/ Lib
DSP Style Programming

C - POSIX-Level
CPU Style Programming

OpenCL
GPU Style Programming

OpenDataPlane
Open API for networking

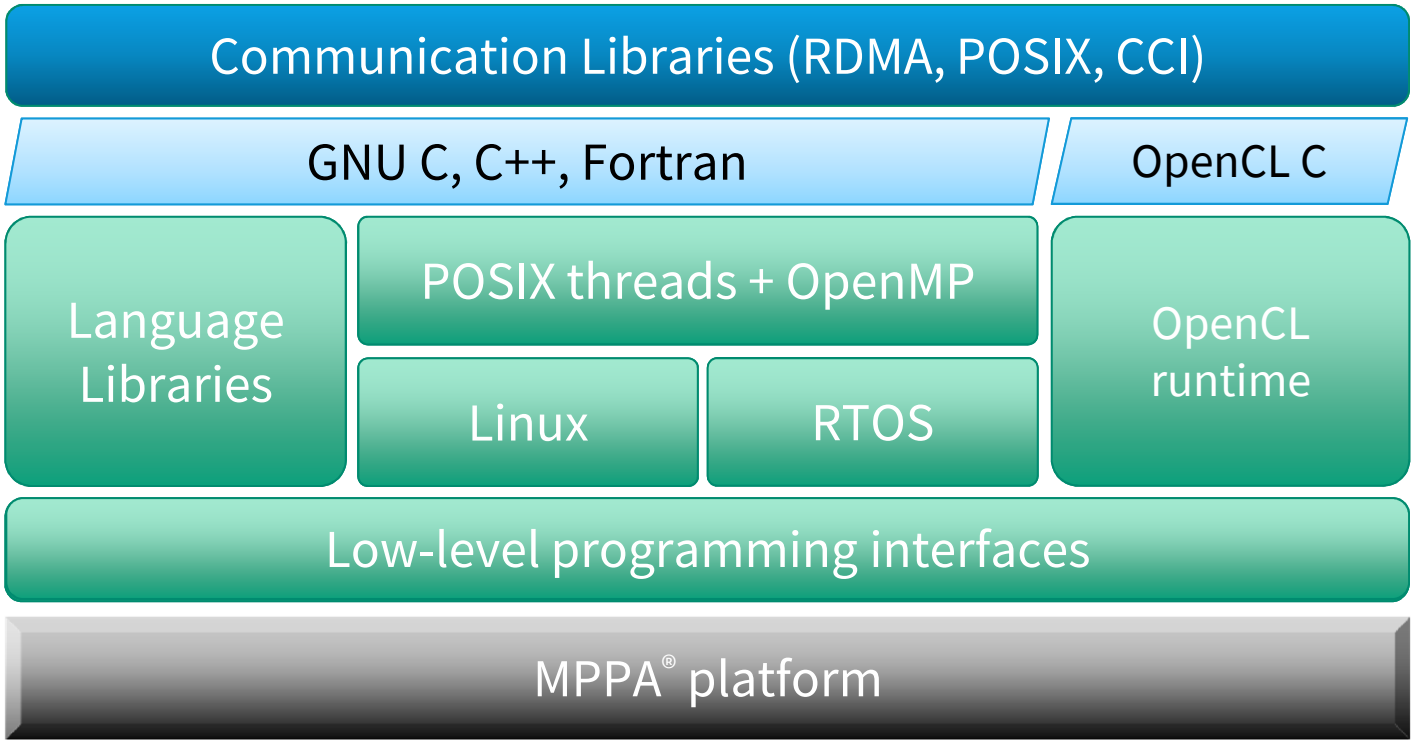
Linux

eSOL

ANSYS

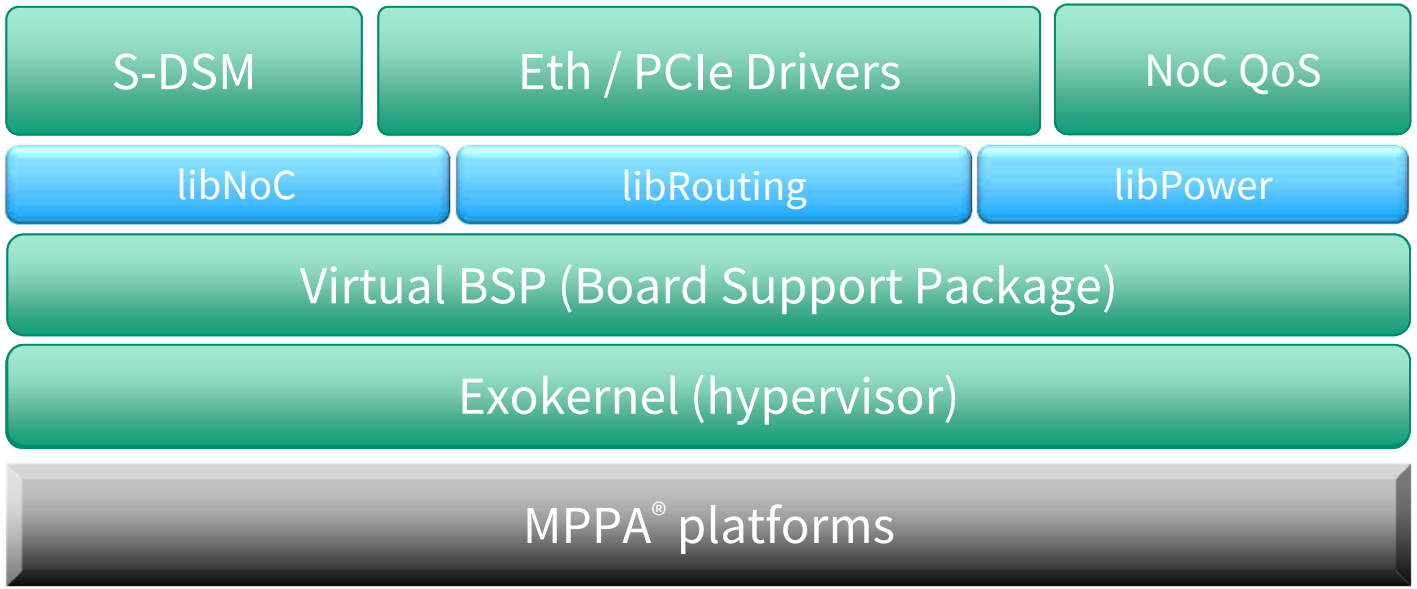
ESTEREL technologies

MPPA® Software Stack Overview



Low-Level Programming Environment

- Performance programming and 3rd party RTOS interface
  - Exokernel + Virtual-BSP + LibNoC, LibRouting, LibPower
  - Support of software DSM (Distributed Shared Memory)
  - Implement the Low-Level programming interfaces



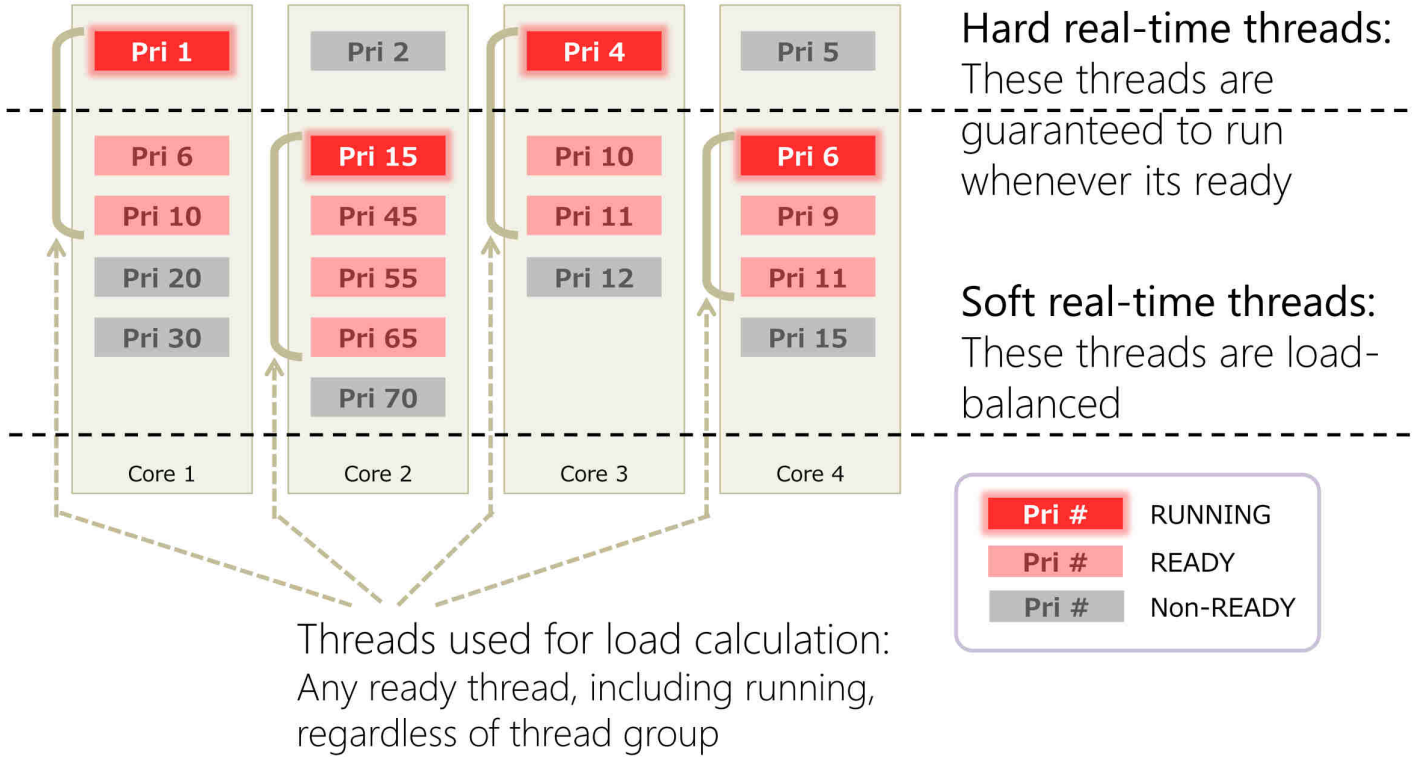
Operating Systems & Device Drivers

- NodeOS for compute clusters
  - Provides POSIX threads, timers and run-time support for GCC OpenMP
  - The RM manages the NoC interfaces and supports the security/safety functions
  - The PEs execute application code on top of exokernel, one thread per core
- SMP Linux on I/O clusters
  - Running on a Kalray quad-core, the other quad-core manages the clusters
  - Device drivers for flash, I2C & SPI (sensors, small peripherals), GPIO
  - OpenMP, uClibc C library, gcc –fdpic
  - PREMPT-RT patch for real-time
- 3rd-party RTOS and middleware
  - eSOL eMCOS is the world's first commercially available manycore RTOS for use in embedded systems first ported to Tilera, now supported on the MPPA® processors
  - ERIKA Enterprise, the first open-source OSEK/VDX (automotive) certified RTOS

eMCOS for MPPA® Processors

Embedded Excellence™

- eSOL eMCOS is the world's first many-core RTOS for embedded use
  - Runs on the MPPA® clusters on top of Low-Level programming



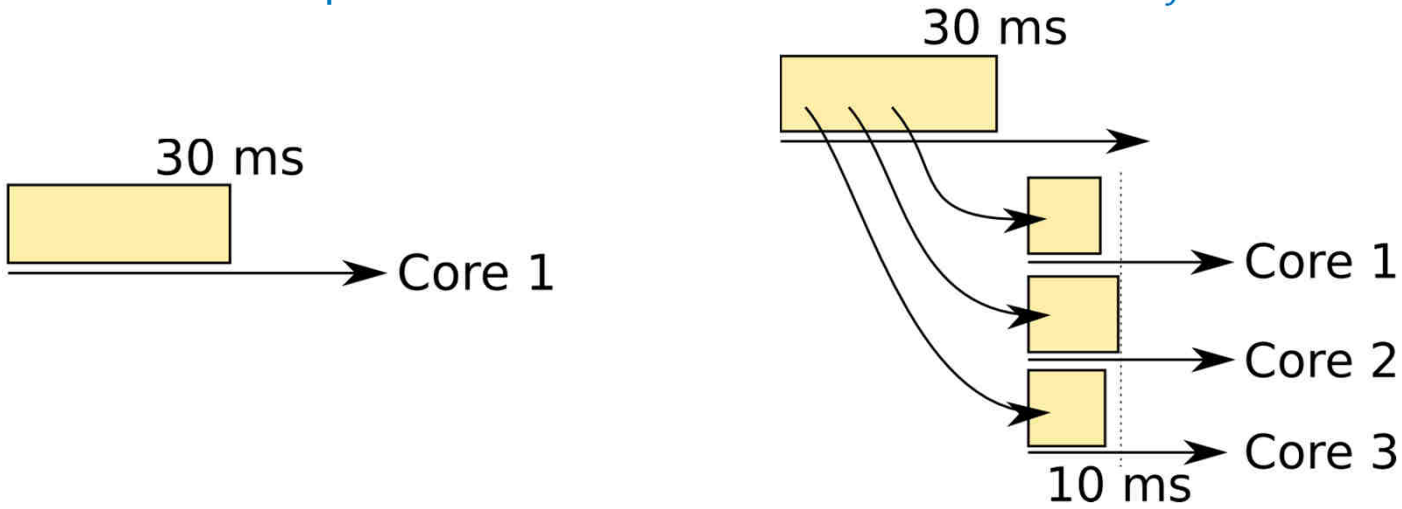
Applications of MPPA® MANYCORE Processors

- Cloud and Data Center acceleration
  - Offloading of real-time or compute intensive functions from x86 applications
  - Domains of application: video, networking, storage, OHPC, data analytics, cybersecurity
  - MPPA® Compute Clusters seen as OpenCL Compute Units or pools of DSP processors

- High Performance Embedded Computing
  - Stand-alone computing enables increased integration of functions including those constrained by real-time
  - Domains of application: aerospace, automotive, transport, energy
  - MPPA® Compute Clusters seen as precision-timed multicore CPUs

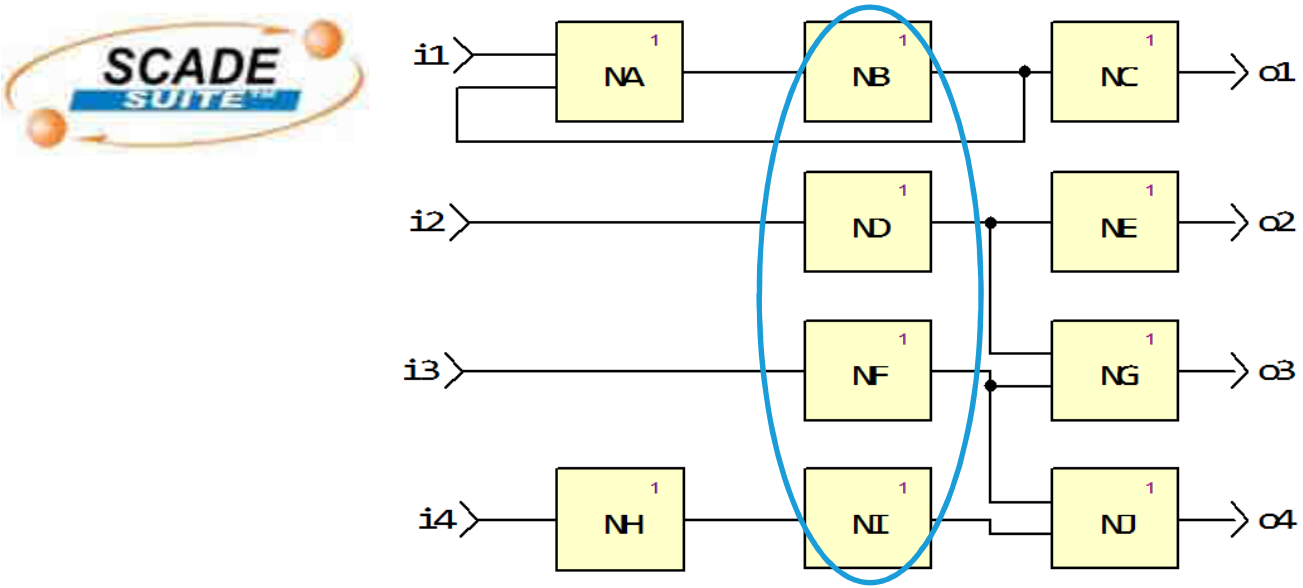
SCADE Code Generation for the MPPA®

- Safety-critical control-command applications
  - Model-based programming using SCADE Suite® from Esterel Technologies
  - Complemented with static timing analysis of binary code (aiT from AbsInt)
- Motivations for multicore and manycore execution
  - Distribute the compute load across cores and reduce memory interferences



  - Effective implementation of multi-rate harmonic real-time applications
  - Envision use of fast Model Predictive Control (MPC) techniques

Example of SCADE Program with Annotations



- Original SCADE program

```
b = B(a);
d = D(i2);
f = F(i3);
i = I(h);
```

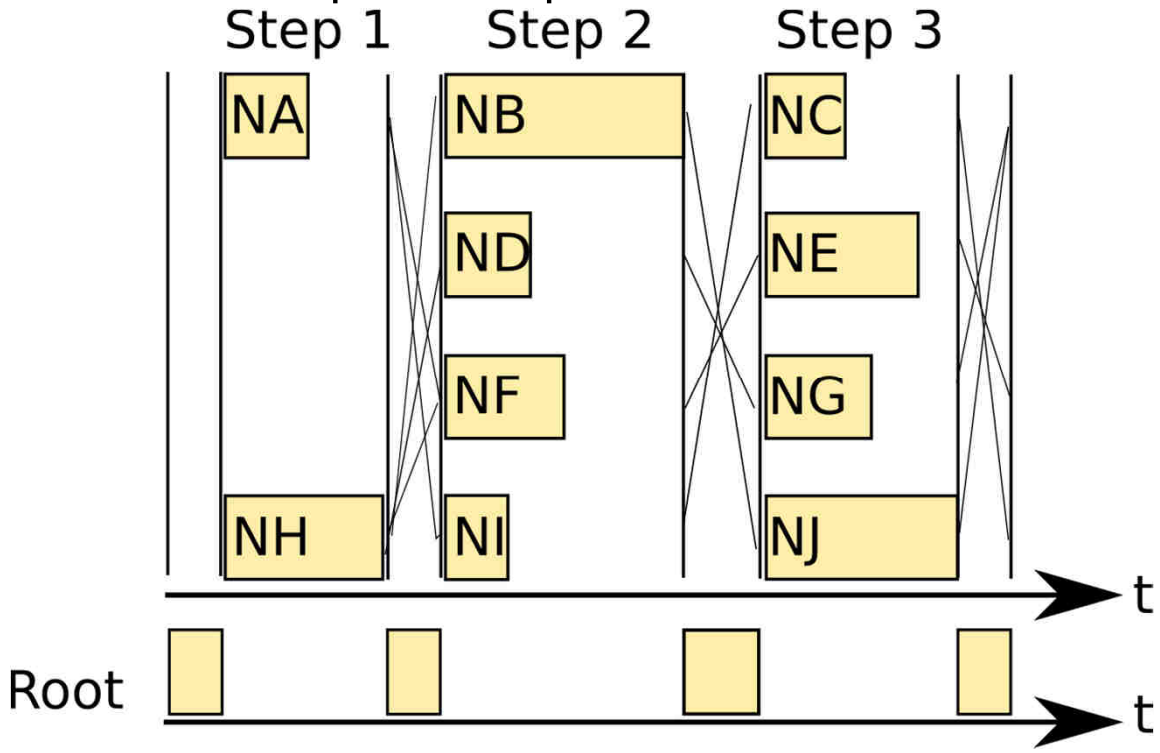
- Annotated SCADE program

```
b = #par_2 B(a);
d = #par_2 D(i2);
f = #par_2 F(i3);
i = #par_2 I(h);
```

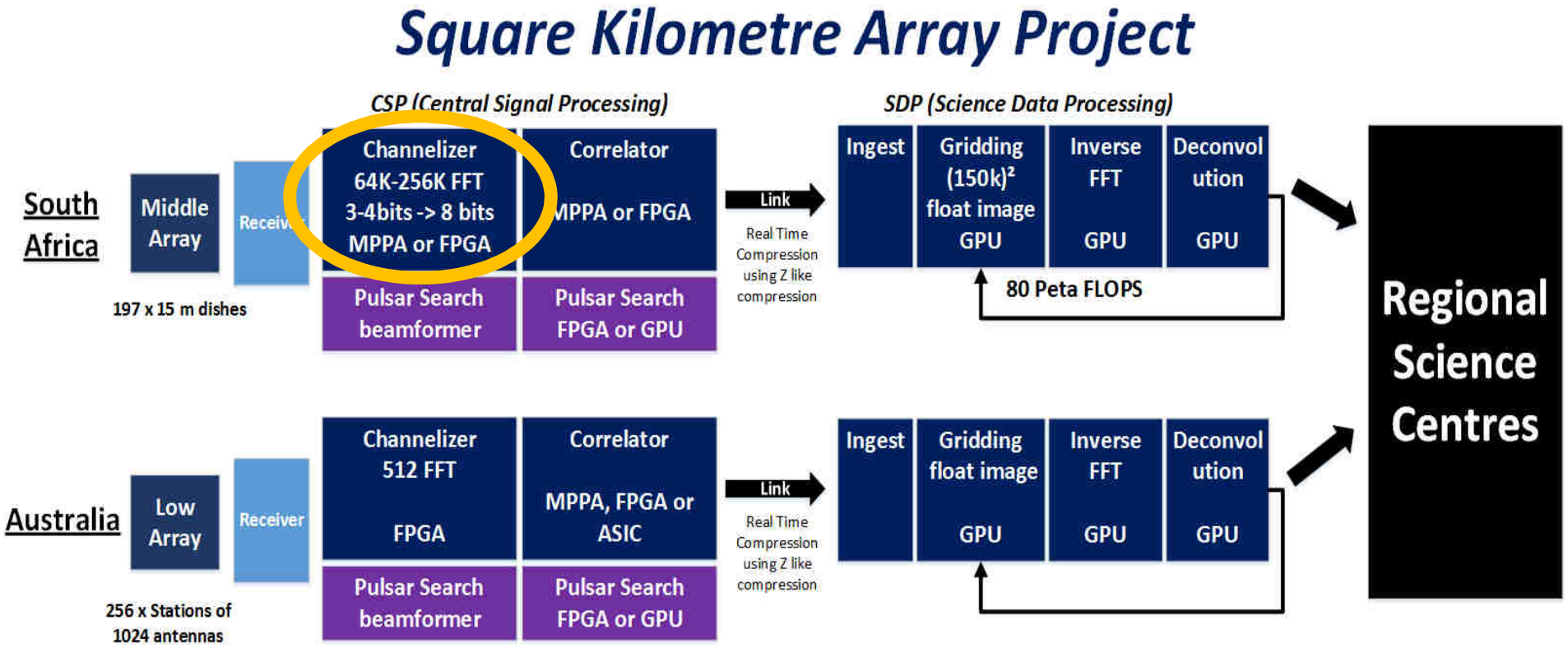
The Bulk Synchronous Parallel (BSP) Model

The BSP is a ‘bridging model’ for parallel computing proposed by L. Valiant

- Program executes as a sequence of ‘supersteps’
- No communication effects until after a ‘superstep’
- This is how we interpret the #par annotations in SCADE

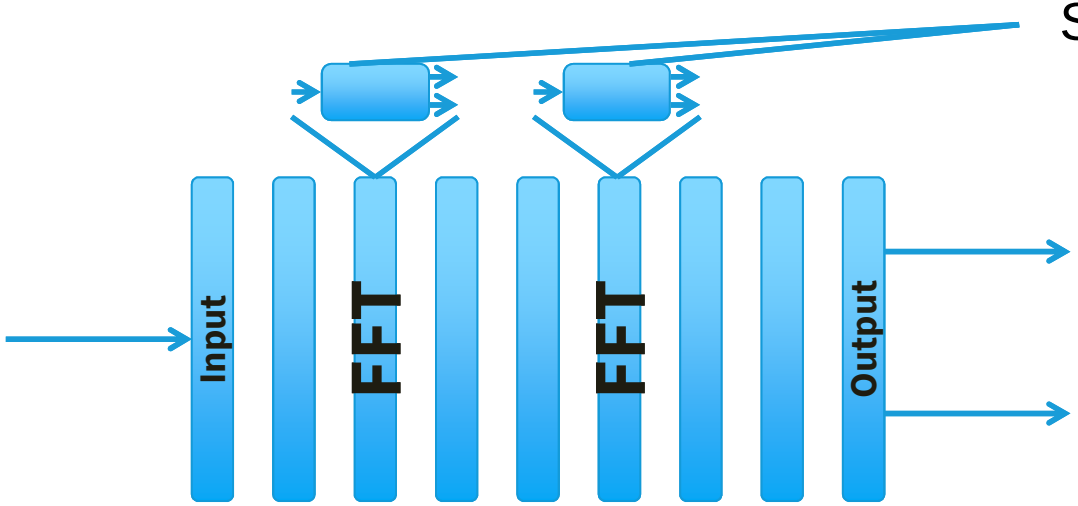


Real-Time Signal Processing with the MPPA®



The 6-step FFT Implementation on the MPPA®

1. Transposition of matrix $N * M$
2. Compute M times N - point - FFT
3. Transposition of matrix $N * M$
4. Twiddle Correction $Mat_{out}(x, y) = Mat_{in}(x, y) * e^{-\frac{2i\pi xy}{N*M}}$
5. Compute N times M - point - FFT
6. Transposition of matrix $N * M$



Simple 512-point FFT Radix-2

16 cores per compute cluster
32 512-point FFT per core
(M=16*32 N=512)

SKA Channelizer Performances on the MPPA®

Implementation*	MPPA (Andey) Compute Cluster Timing (ms)	MPPA-256 Andey @ 400MHz (ms)	MPPA-256 Bostan @ 600MHz (ms)	Accuracy**
Reference implementation : Fully Fixed-point – 6 steps (2014 Implementation)	11,64	**1,07**	***0,65***	+/- 60 integer
Mix Float Implementation : Floating-point in FFT stages with fixed-point storage in the middle of the 6-step	14,73	1,26	*0,70*	+/-3 integer

**: All including communications in the performance measurement: combine and samples arrangement*
***: Precision difference on the worst samples WITHOUT any scaling Input test gaussian mean 0 std 9*

Running Multiple ADAS-Oriented Functions on a Single MPPA®-256 Processor (Demonstration)

DEMO System

Software

Sobel Filter

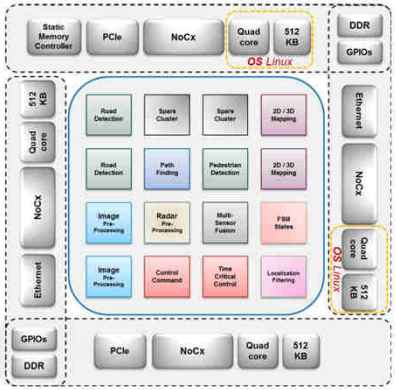
Edge detection

Hough Transform

Line detection

Harris Filter

Feature points detection



MPPA®-256 PCIe Application Board AB01

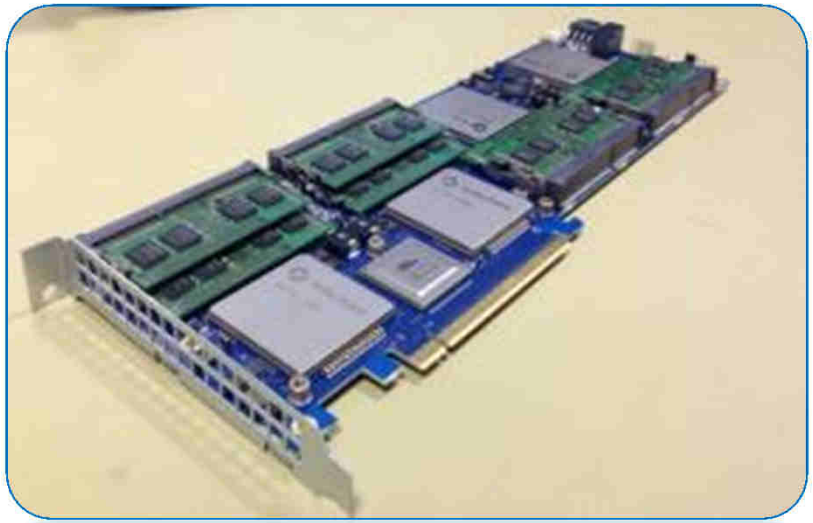
- Interfaces to the 2 I/O clusters
  - 2 PCIe GEN3 x8 interfaces through a x16 PCIe switch
  - 2 DDR3 interfaces
  - 4 Ethernet interfaces (2x10G + 2x1G)
  - 1 NoCX interface
  - NOR flash, GPIOs, leds, buttons, extensions & debug connectors

PCIe Cards for the MPPA®-256 Bostan Processor

KONIC-80

- 1x MPPA®-256 Bostan processor
- 2x QSFP+ => 2 x 40GbE or 8 x 10GbE
- 2x DDR3 @ 2133 MT/s => 34 GB/s
- OpenDataPlane SDK
- Virtualization Offload
- First engineering samples: Q4-15
- Volume Production: Q1-16

TurboCard-3



- 4x MPPA®-256 Bostan processors
- 2x NoC Extension interfaces
- 2.5 TFLOPS SP / 1.25 TFLOPS DP
- 8x DDR3 @ 2133 MT/s => 136 GB/s
- OpenCL SDK
- First engineering samples: Q1-16
- Volume Production: Q2-16

Thank you

KALRAY S.A.
Paris - France

86 rue de Paris,
91 400 Orsay
France

Tel: +33 (0) 184 00 00 45
email: info@kalray.eu

KALRAY S.A.
Grenoble - France

445 rue Lavoisier,
38 330 Montbonnot
France

Tel: +33 (0)4 76 18 09 18
email: info@kalray.eu

KALRAY INC.
Los Altos - USA

4962 El Camino Real
Los Altos, CA
USA

Tel: +1 (650) 469 3729
email: info@kalrayinc.com