

# Update on Ara

12/10/2022

**Matteo Perotti**

**Matheus Cavalcante**

**Nils Wistoff**

**Professor Luca Benini**

**Integrated Systems Laboratory**

**ETH Zürich**

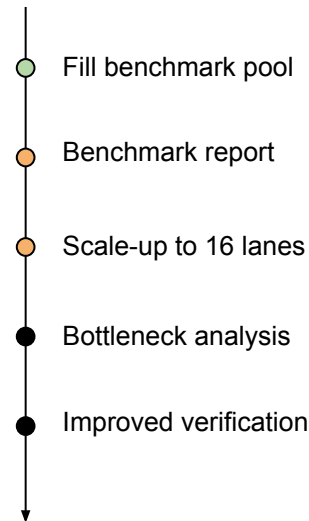
# Summary

- **Software**

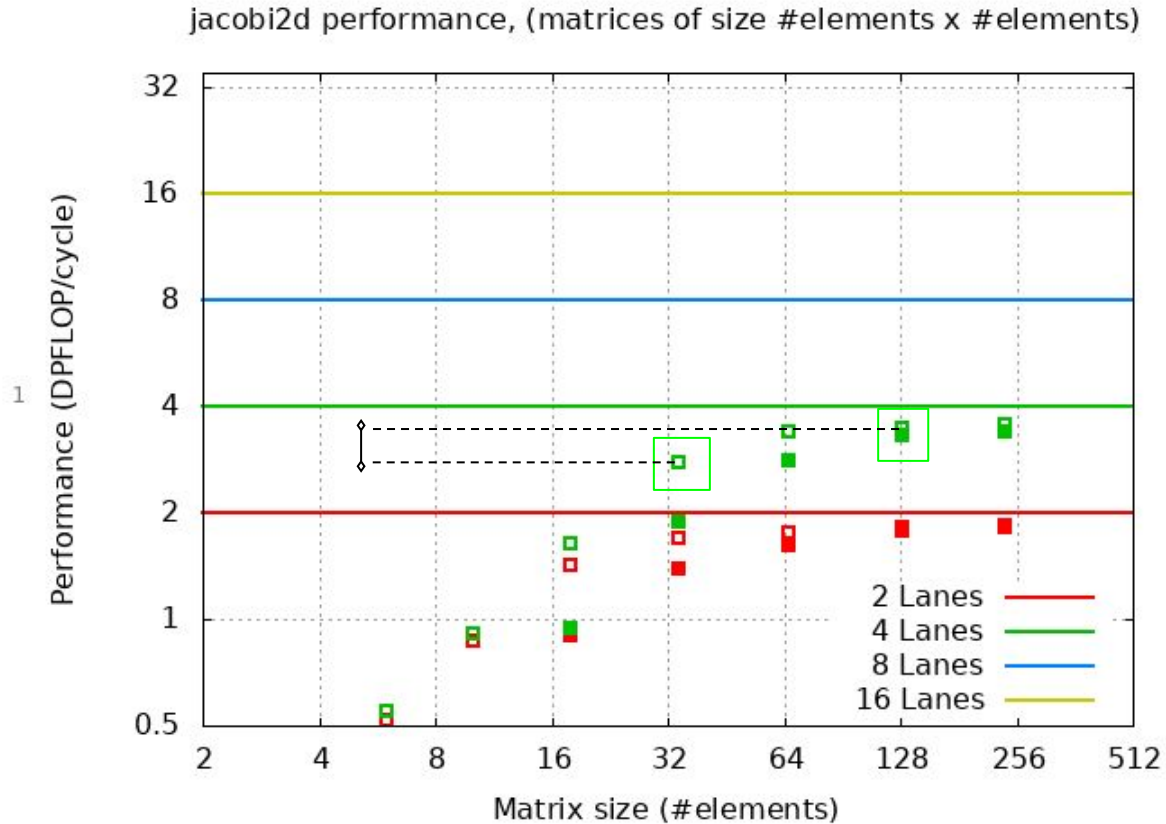
- Analyze kernels
- Try (last time) CI ideal dispatcher

- **Hardware (RTL + Backend)**

- Scale to 16 lanes
- Merge Fixed-Point support



# Jacobi2d

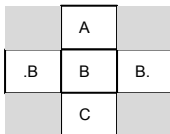
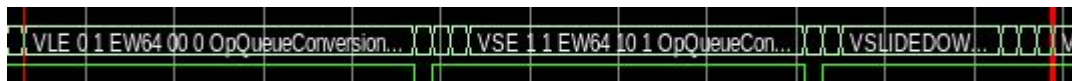


## Jacobi2d

vfadd, vfmul

vslide1up

3 x vfadd



```

vld      v0, @i
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o

```

```

vld      v0, @i
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1

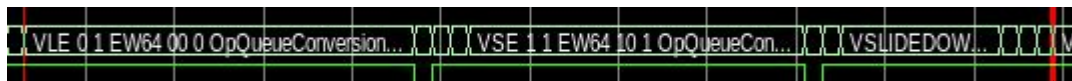
```

## Jacobi2d

vfadd, vfmul

vslide1up

3 x vfadd



VSTORE (VLSU)

Load cannot fire during a store and vice-versa!

vfadd, vfmul cannot reach Ara's FPU on time

```

vld      v0, @i
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1

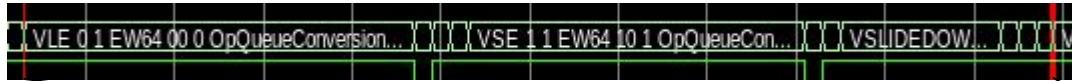
```

## Jacobi2d

vfadd, vfmul

vslide1up

3 x vfadd



VLOAD (VLSU)

Load cannot fire during a store and vice-versa!

vfadd, vfadd, vfadd cannot reach Ara's FPU on time

```

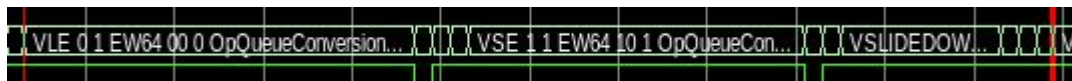
vld      v0, @i
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1
  
```

# Jacobi2d

vfadd, vfmul

vslide1up

3 x vfadd



## Back2Back FPU dependant operations

**RAW** hazard - More problematic with short vectors!

The first fadd finishes issuing the operands before writing back a result

Second fadd stalls two cycles

Reorganize fadd? Earlier dependency with VLSU!

```

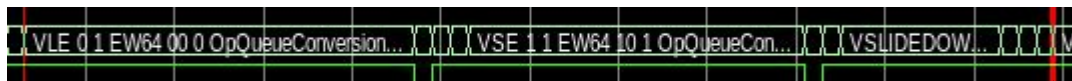
vld      v0, @i
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1
  
```

## Jacobi2d

vfadd, vfmul

vslide1up

3 x vfadd



vslide1down still cannot chain!

The **RAW** hazard is shifted within the vector

```

vld      v0, @i
fadd     v8, v8, v2
fmul     v8, v8, n
vst      v8, @o
vslide1up v1, v0
vslide1down v2, v0
fadd     v8, v3, v4
fadd     v8, v8, v0
fadd     v8, v8, v1
  
```

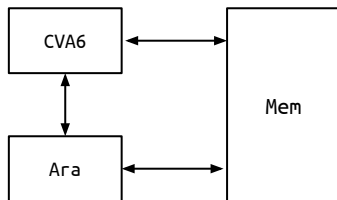


# PolAra

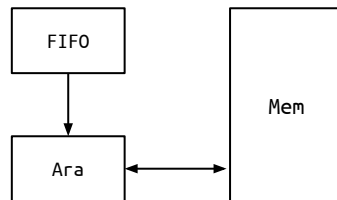
- **Ara TO in gf22**
- **Cooperation with another University**
- **Keep it simple, then add feature**
  - **Start from Yun SoC**

# Software - Ideal Dispatcher

Default System



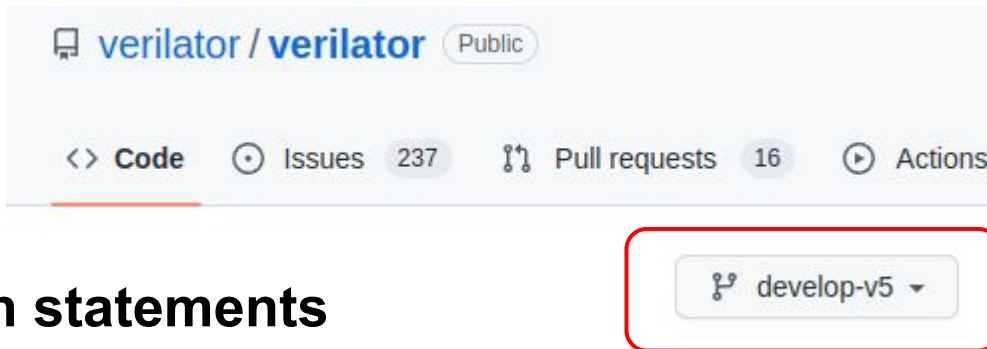
Ideal Dispatcher System



**Ideal issue rate from FIFO to Ara**

# Software - Ideal Dispatcher


- No verilator?
- New beta version **V5**
  - Supports non-synth statements
- Still **issues** with
  - **wait()** statements
  - **Trace**



# RVV1.0 Compliance

- **Cooperation with Company**
- **Add missing instructions**
  - Still **iterating** through the flow (Ara gets larger)

☐  [HW] Add support for vector mask instructions

#149 opened 4 days ago by M-ljaz-10x  3 tasks done

☐  [HW] Add support for vector fixed-point instructions

#147 opened 14 days ago by M-ljaz-10x  3 tasks done

# Projects

- **Verification**
  - Verify and Fix Ara
  - Force RISC-V + Compliance
  - Supervisor from company
- **OS Support**
  - Virtual memory
  - Interrupts
  - Verification

# Backend

- **16 Lanes flat flow** (very slow...)
- Ara 8 Lanes new floorplan
- Fusion flow - Some debug needed
- Still ongoing!

# Verification

- During compliance verification...
- **Reshuffle bug!**
  - **Critical for compliance**
- **It should be fixed now**



Commits on Oct 11, 2022

[hardware] 🚀 Reshuffle source vregs for in-lane ops ...

In case of an operation executed within the lanes (e.g., vadd), the lanes cannot read the vector correctly if the vseq of the source register is different from the eew\_q of that register (different byte layout!). So, we need to reshuffle in this case as well. Since this is costly, the suggested software procedure is to avoid filling a register with an encoding and using it with a different one.

μ mp-17 committed 14 hours ago

[hardware] 🚀 Fix wrong assignments in operand queues

μ mp-17 committed 14 hours ago

[hardware] Clarify eew-related signals

μ mp-17 committed 14 hours ago ✓

# Further

- **Software**

- Merge ideal dispatcher branch
- Gather performance data + Report

- **Hardware (RTL + Backend)**

- Try different die shapes
- Close timing with 16 lanes

