

# Update on Ara

25/01/2021

**Matteo Perotti**

**Matheus Cavalcante**

**Nils Wistoff**

**Professor Luca Benini**

**Integrated Systems Laboratory**

**ETH Zürich**

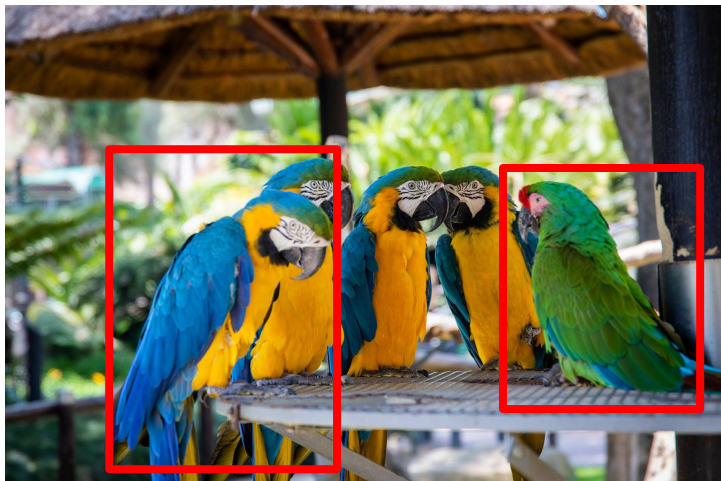
# Summary

- **New benchmark**
  - Rol Align
  - First implementation
  - Bug fixes
- **SDF simulation... towards full power analysis**
  - Hello World
  - fmatmul
- **Projects about Ara**
  - Test Yun
  - Add missing instructions for RVV 1.0

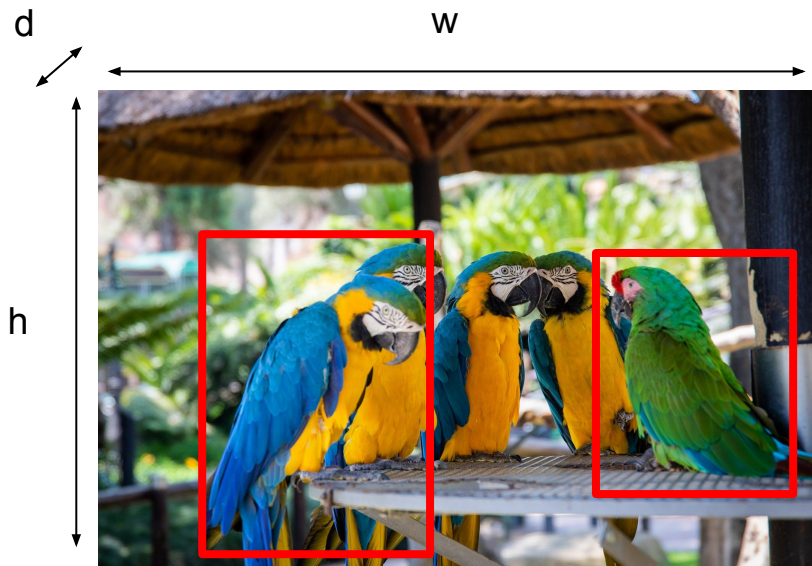
# Rol Pooling and Align



# RoI Pooling and Align

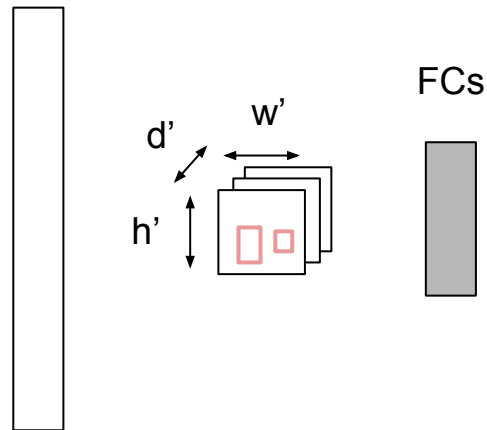


# RoI Pooling and Align



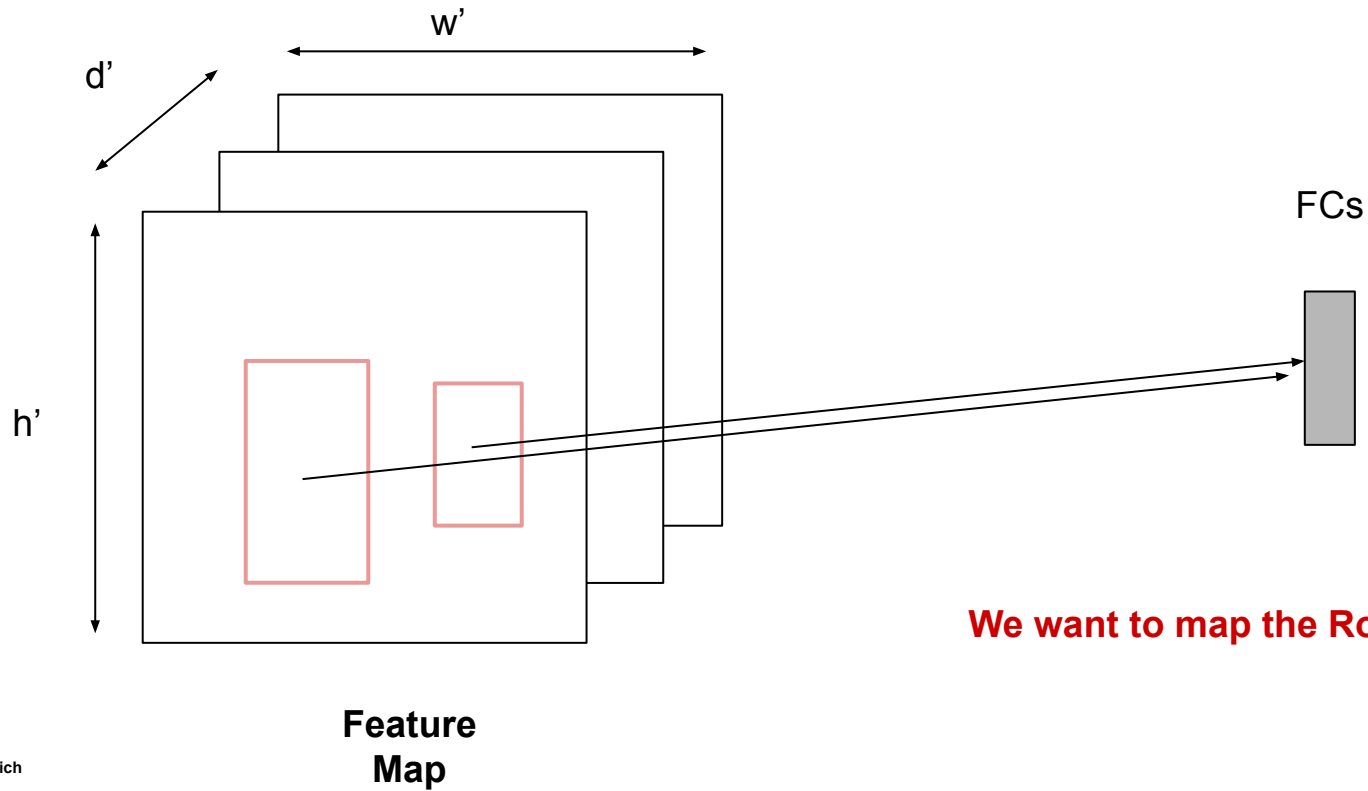
Image

ConvNet

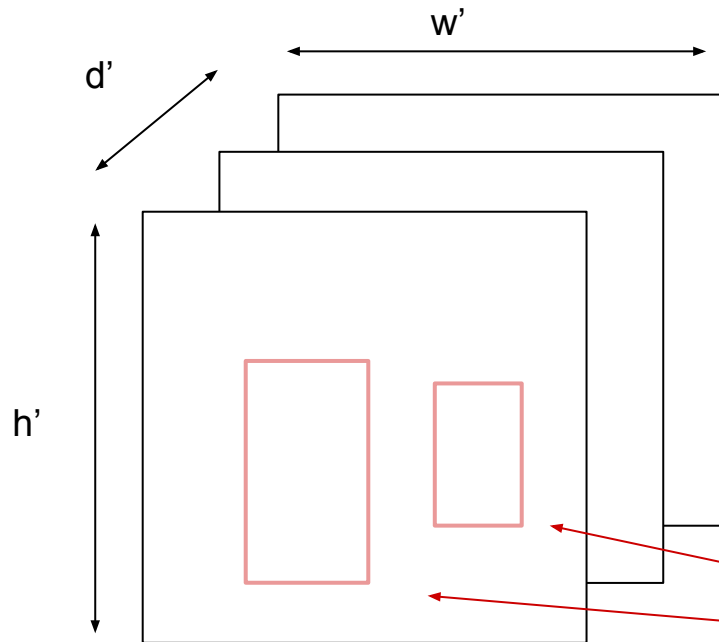


Feature Map

# RoI Pooling and Align

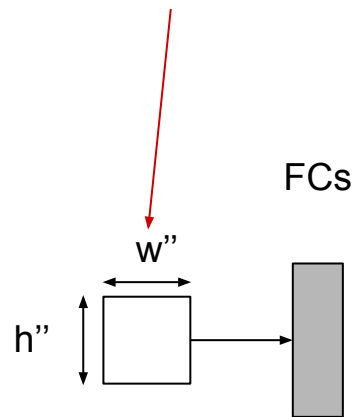


# RoI Pooling and Align



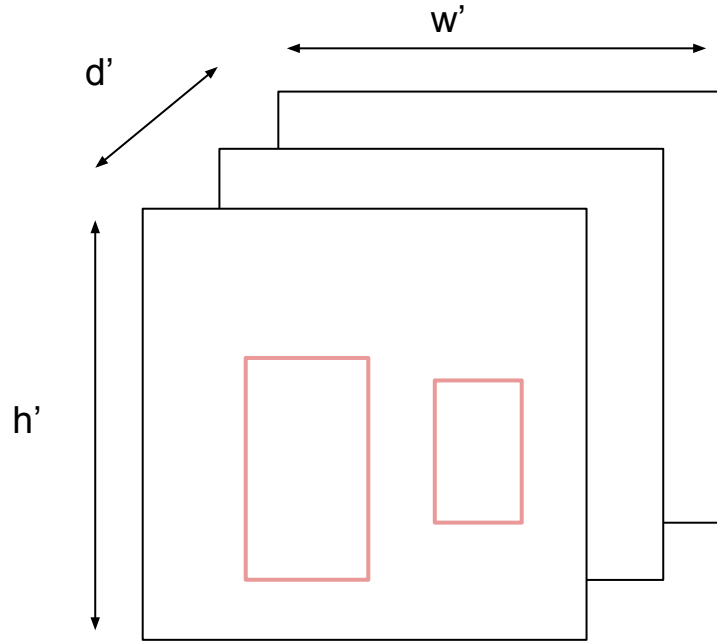
Feature  
Map

Input size of FCs is fixed

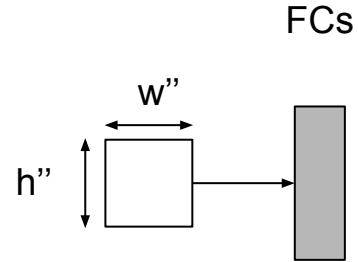


Our Rols have variable size!

# RoI Pooling and Align

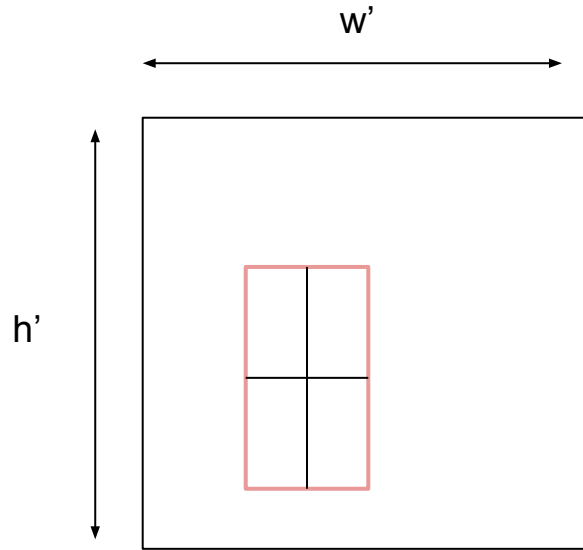


Feature  
Map

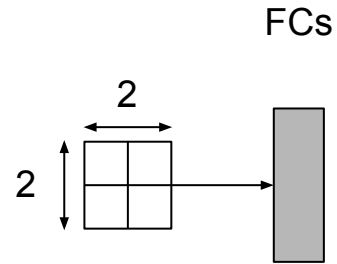




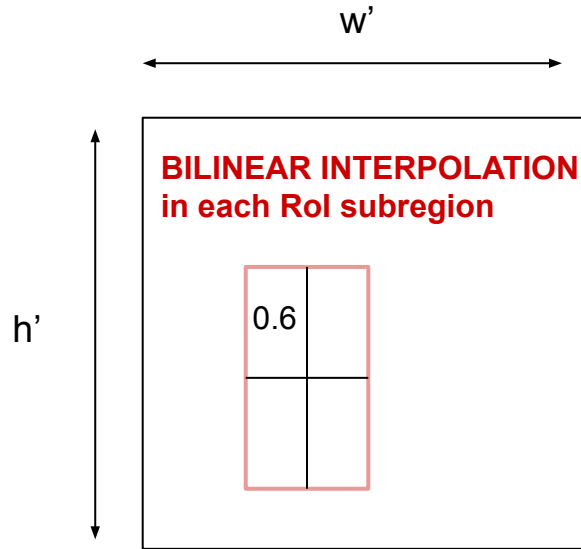
# RoI Pooling and Align



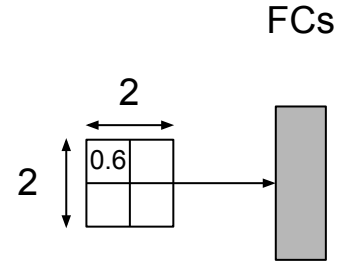
**Feature  
Map**



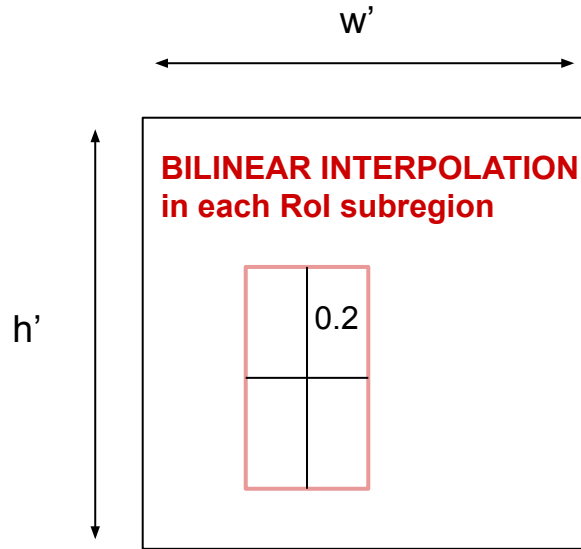
# Rol Pooling and Align



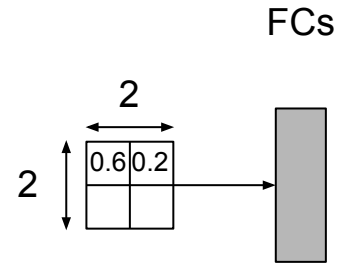
Feature  
Map



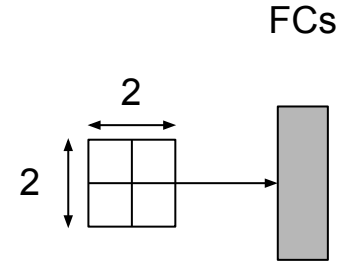
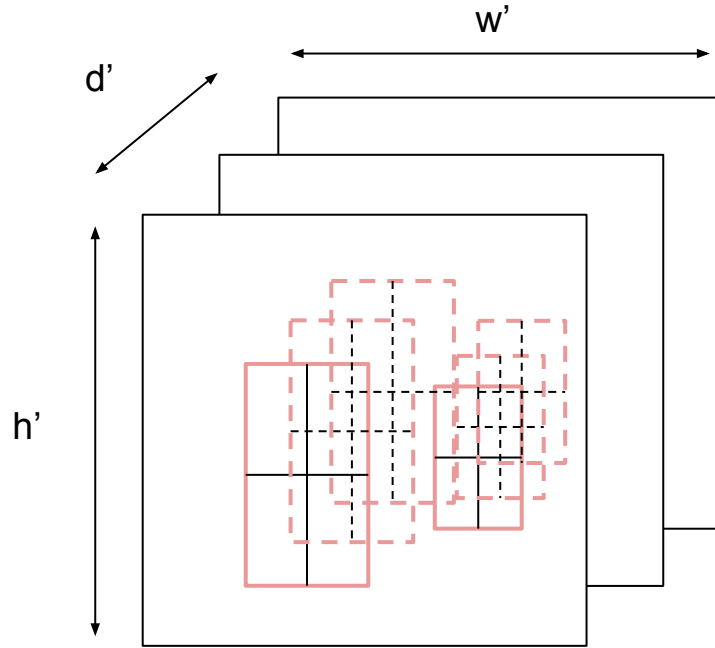
# RoI Pooling and Align



**Feature  
Map**

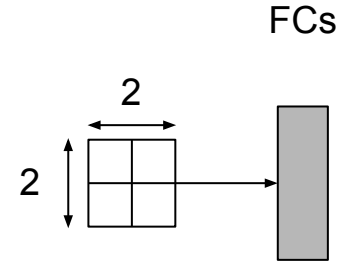
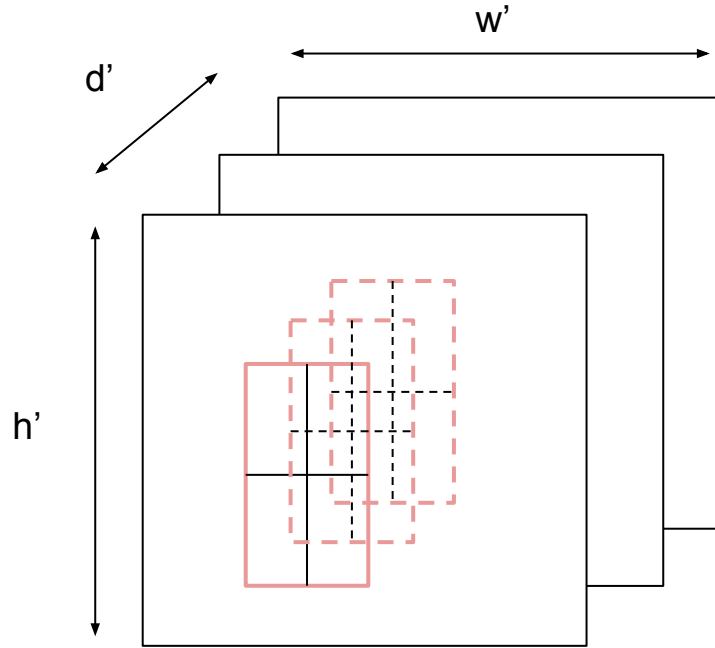


# RoI Pooling and Align



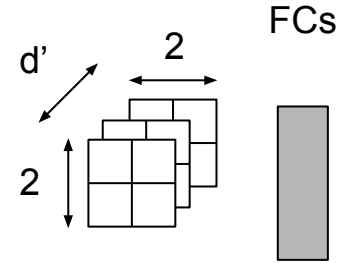
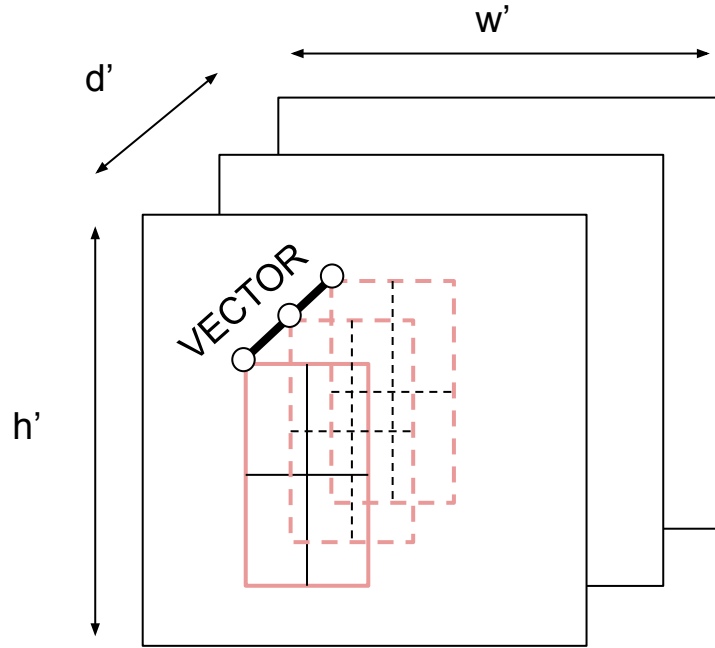
**Repeat for each subregion, for every “channel”, for every RoI**

# Rol Pooling and Align



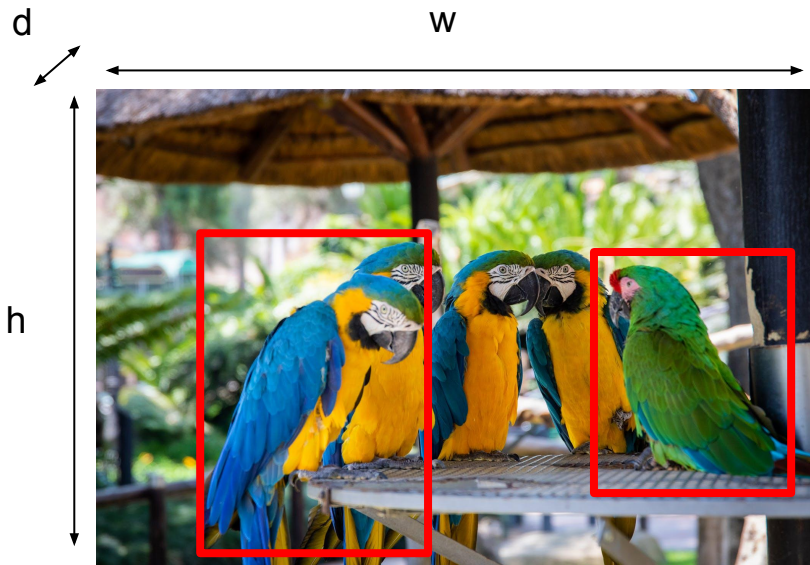
**First Rol Align Vector implementation: parallelize on  $d'$  (depth)**

# RoI Pooling and Align



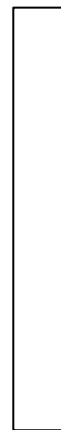
**First RoI Align Vector implementation: parallelize on  $d'$  (depth)**

# RoI Pooling and Align

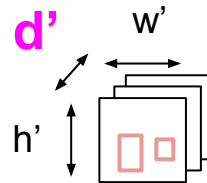


Image

ConvNet



Effective if  $d'$  is high enough



FCs



Feature Map

# Rol Align

- First porting of <https://github.com/longcw/RolAlign.pytorch>



# RoI Align

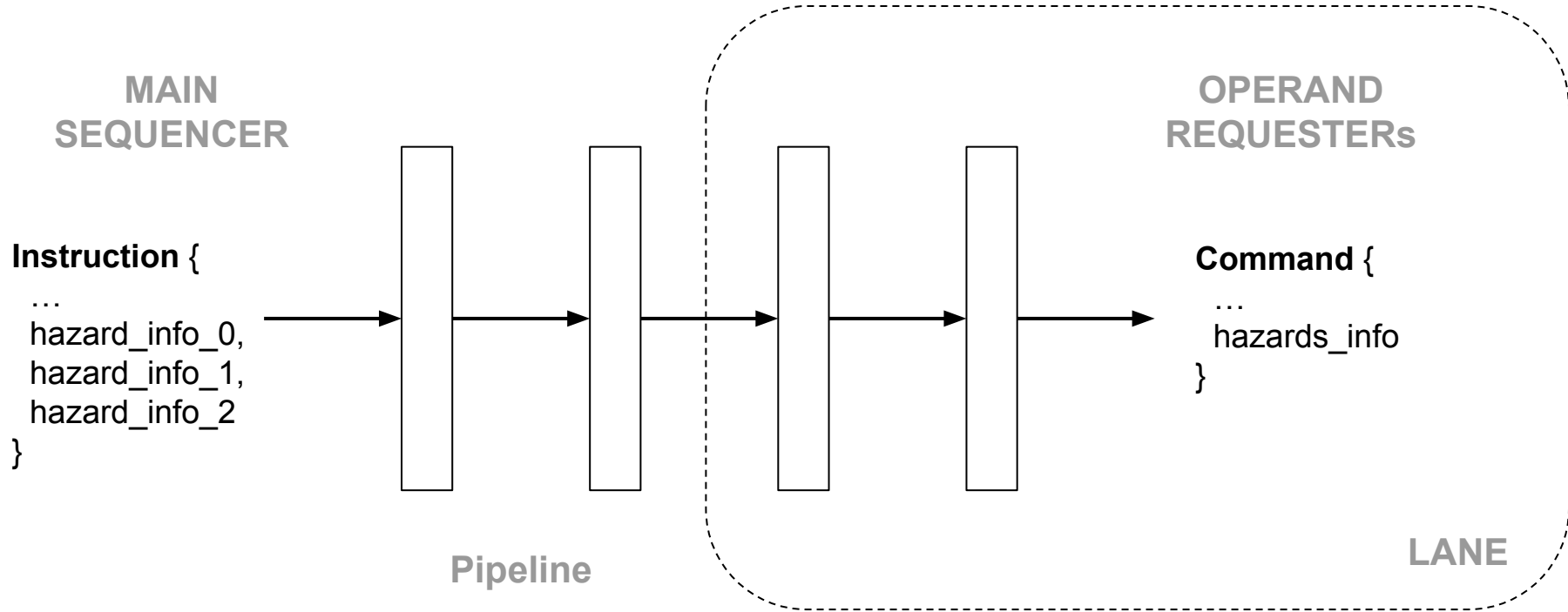
- First porting of <https://github.com/longcw/RoIAlign.pytorch>
- Input feature map ordering in memory?
  - If depth is contiguous: unit-strided memory operations (bursts)
  - Otherwise: strided memory operations (no bursts)

# RoI Align

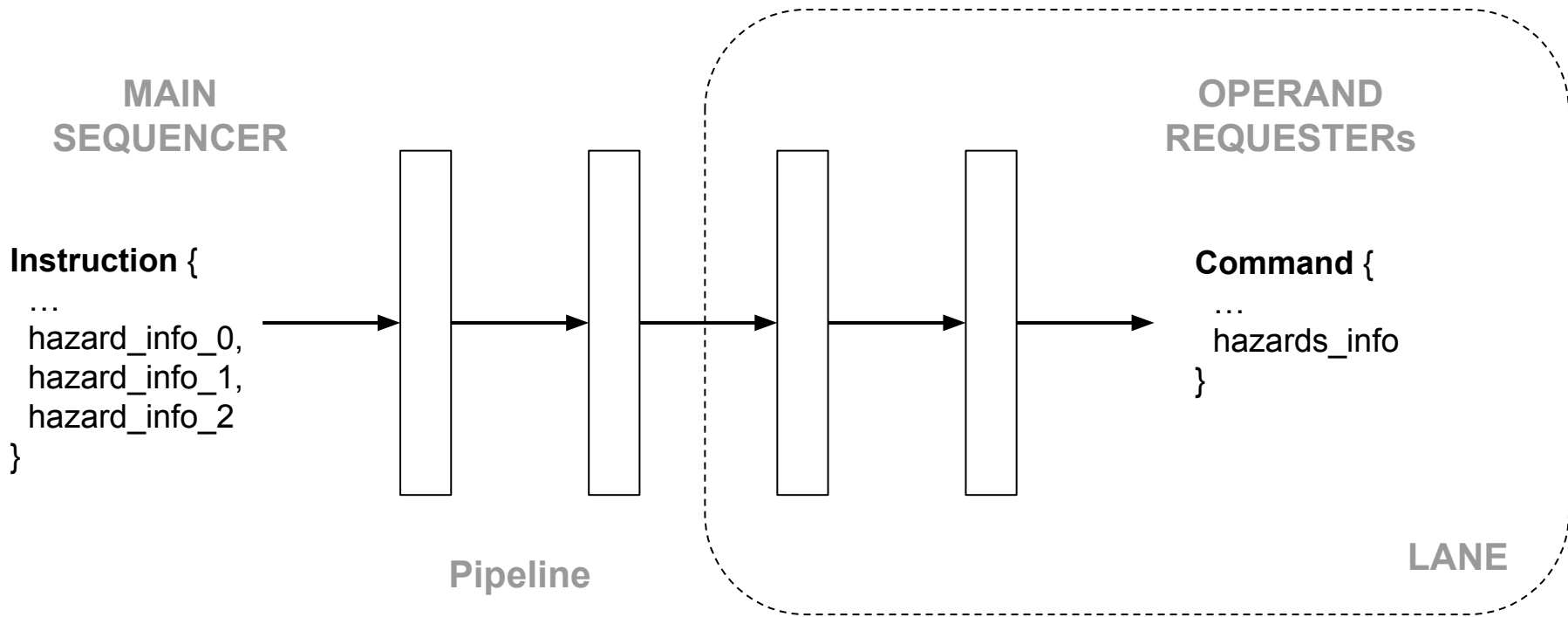
- First porting of <https://github.com/longcw/RoIAlign.pytorch>
- Input feature map ordering in memory?
  - If depth is contiguous: unit-strided memory operations (bursts)
  - Otherwise: strided memory operations (no bursts)
- Bugs on the hazard checks engine

**Bug fixing**

# BUG: dependencies get lost

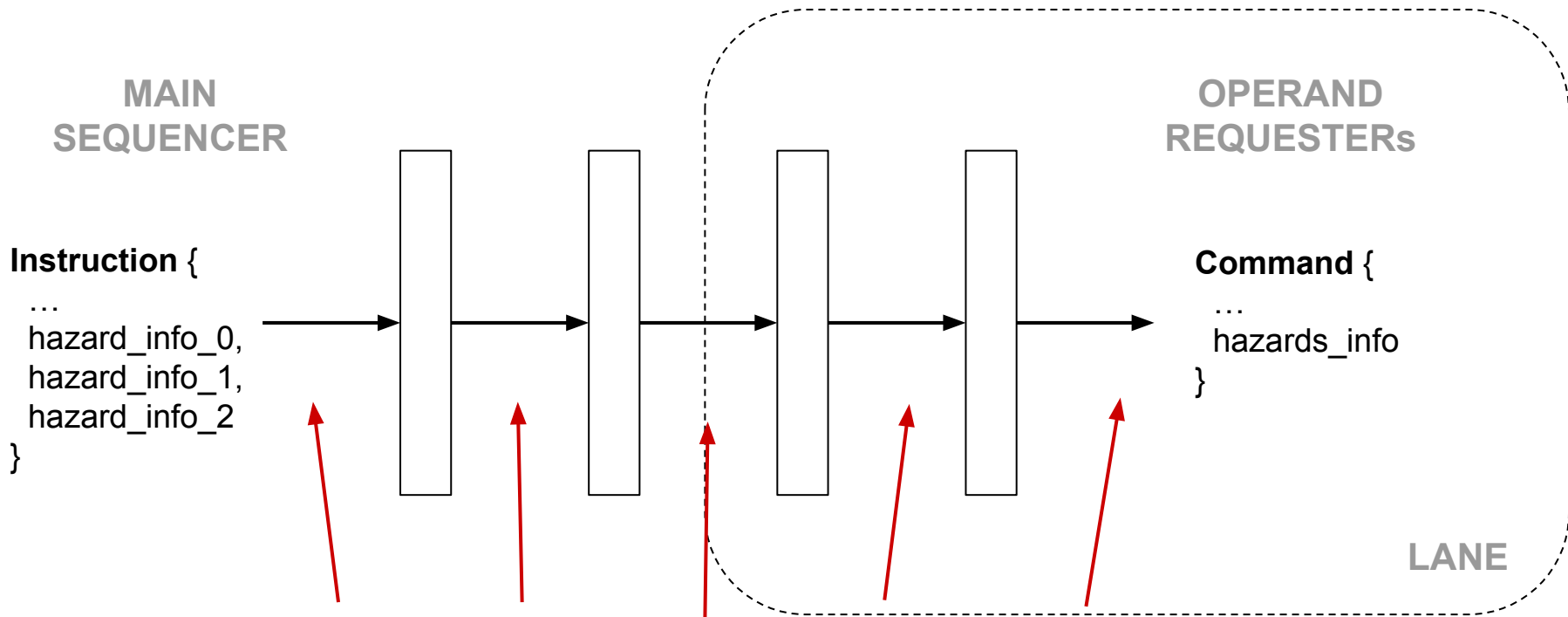


# BUG: dependencies get lost



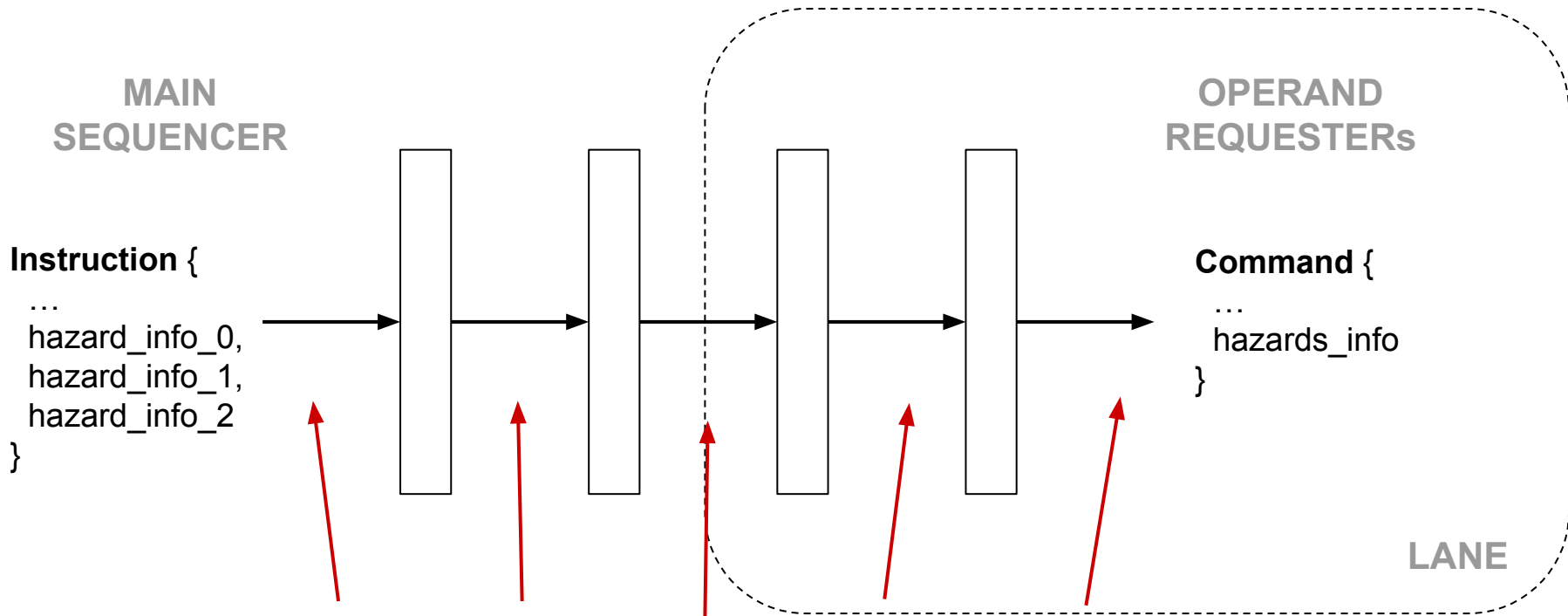
**When a previous instruction finishes, it clears the hazard bits of the dependant instruction**

# BUG: dependencies get lost



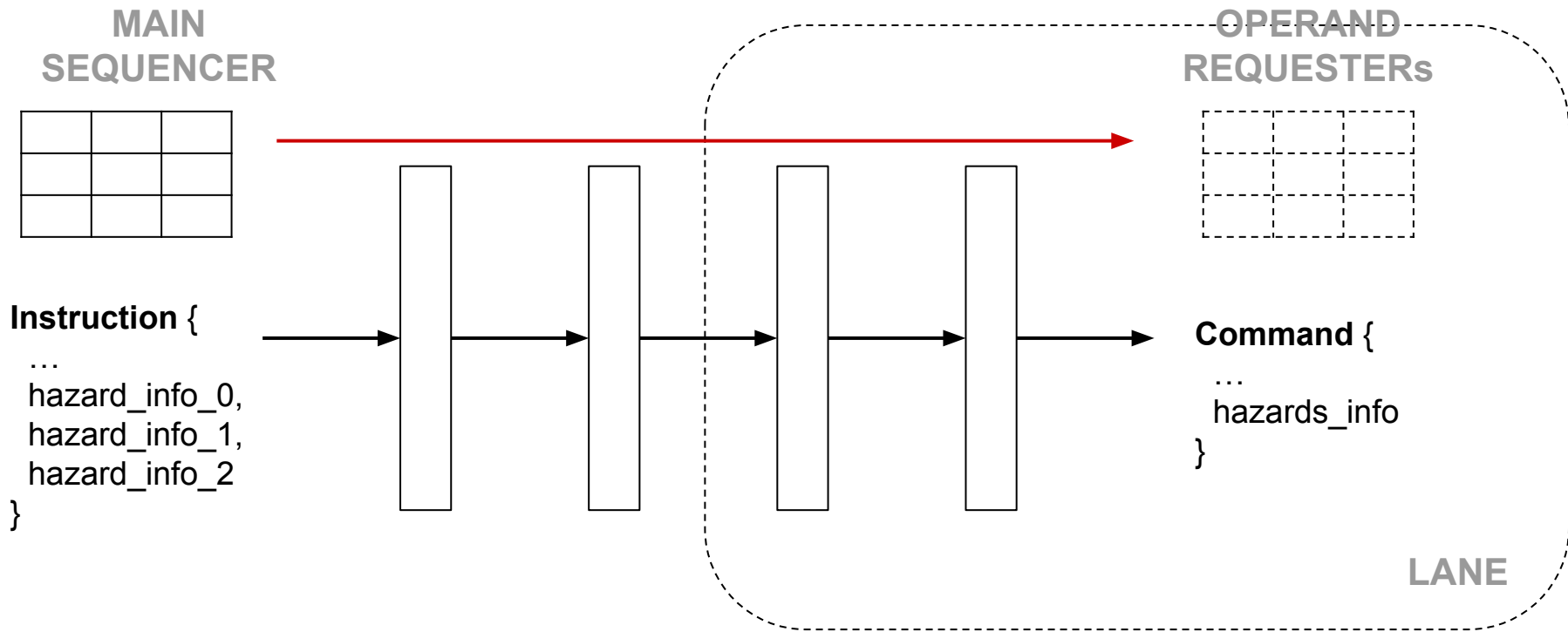
**When a previous instruction finishes, it clears the hazard bits of the dependant instruction**

# BUG: dependencies get lost



**More pipe regs, more places to check for hazard bits clearing**  
**In some parts of the code, this was not easy to do**





# BUG: dependencies get lost



**Solution: GLOBAL HAZARD TABLE in main sequencer**  
**Directly forwarded to the operand requesters**

## Other bugs and misc

- Early grant
  - Some units write in a buffer before the V RegFile
  - Writing to the buffer should not clear the hazard bits
  - Hazard bits cleared when the write-back happens
- FPU fdiv and fsqrt
  - FPU was getting stuck with some fdiv/fsqrt
- Handling PRs of the previous work

<input type="checkbox"/>		7 Open	✓	41 Closed
<input type="checkbox"/>		Hotfix - Main Sequencer	✓	
		#101 opened 6 days ago by mp-17		3 tasks done
<input type="checkbox"/>		Floating-Point classify, division, sqrt	✓	
		#100 opened 7 days ago by mp-17		3 tasks done
<input type="checkbox"/>		Integer Reductions	✓	
		#95 opened 14 days ago by mp-17		3 tasks done



# Simulation with SDF

- Milestone for
  - Verification
  - Accurate Power Analysis
- Hard to make it work with the new flow
- Issues with SDF formats, tools bugs...
- We successfully simulated Hello World!
- We are working on the fmatmul

# Projects

- Yun is back
  - Test it and measure power consumption and energy efficiency
- Student Projects
  - Extend Ara with the missing features for RVV 1.0