Update on Ara

22/06/2022

Matteo Perotti
Matheus Cavalcante
Nils Wistoff
Gianmarco Ottavi

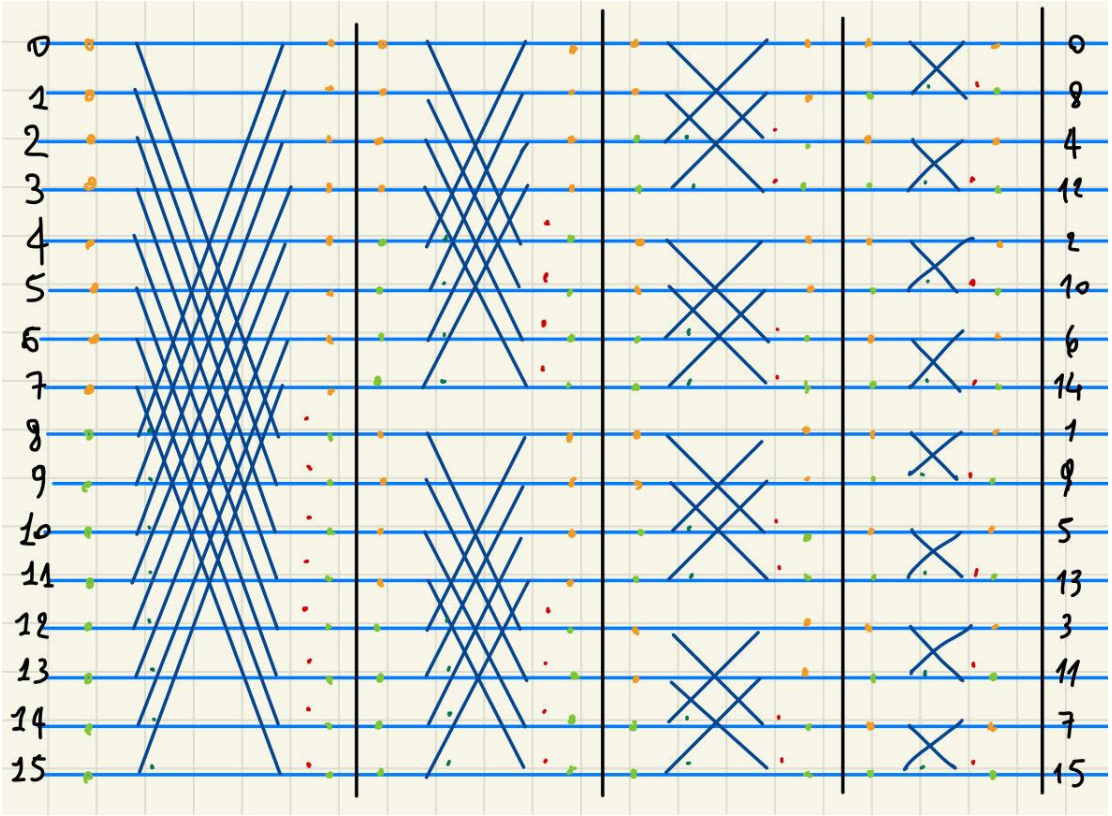
Professor Luca Benini
Integrated Systems Laboratory
ETH Zürich

FFT timeline

- ✓ Python golden model
- ✓ Scalar DIT + DIF (CVA6 only)
- ✓ Vectorized DIF algorithm
  - ✓ First complete implementation
  - ✓ Debugging
- ➢ Performance analysis
- ✗ Optimization
- ✗ ISA extension?

DIF FFT - Status

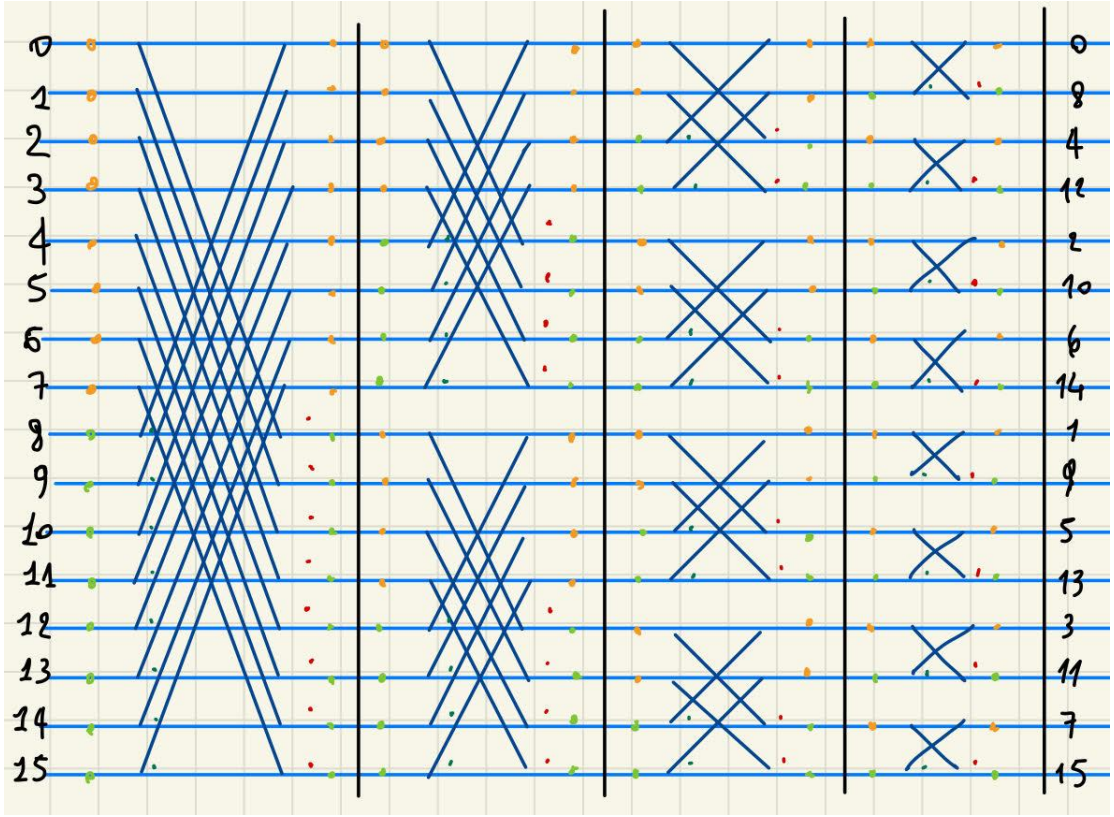
✓ **Ara (4 lanes) works**



DIF FFT - Status

- ✓ Ara (2, 4, 8, 16 lanes) works with up to 256 samples

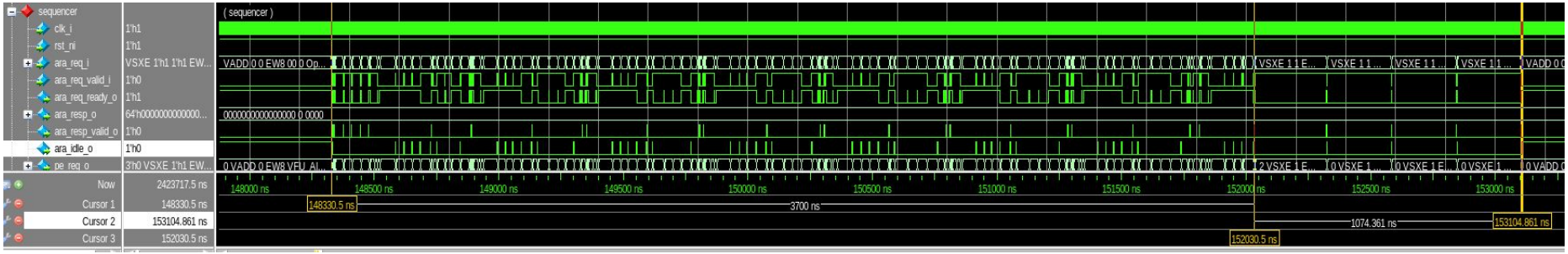
- In memory:
  - Twiddle factors
  - Mask vectors for permutations
  - Indexes for final scatter



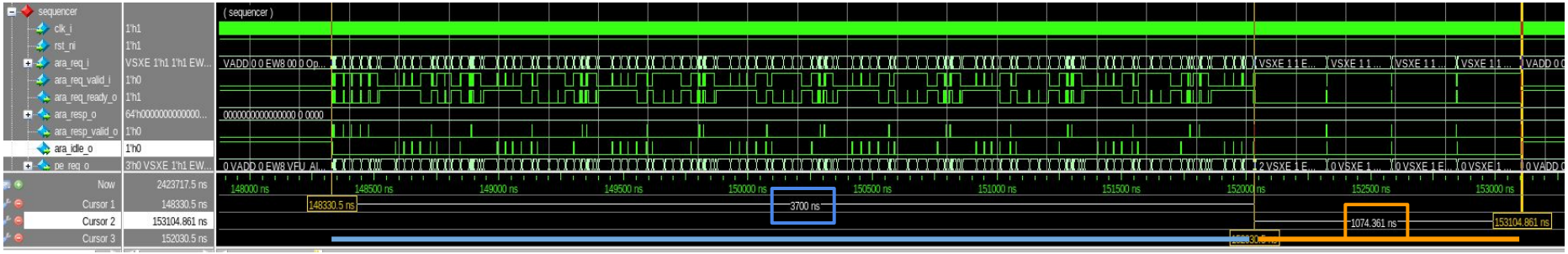
Many exposed bugs!

- First benchmark that stresses all the units together
- Fixed bugs in:
  - Dispatcher
  - Slide unit
  - Mask unit
- Increasing problem size with 16 lanes (more memory!)

Vector DIF FFT (4 Lanes, 256 samples)



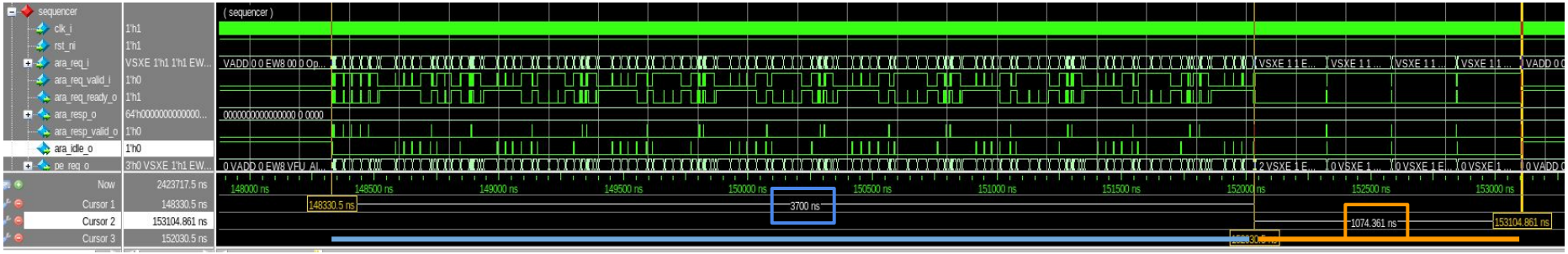
Vector DIF FFT (4 Lanes, 256 samples)



FFT
~3700
cycles

Scatter
~1000
cycles

Vector DIF FFT (4 Lanes, 256 samples)



FFT
~3700
cycles

Scatter
~1000
cycles

- Vector DIF FFT
  - 4700 cycles
- Scalar DIF FFT:
  - 35100 cycles

7.5x performance improvement

Tend to 9x with unit-strided stores

Timeline

- FFT
- ✓ Memcpy benchmark
- FP reduction and MASKU instruction integration
- VFIRST instruction
- Strncpy, strncmp benchmarks
- DWT, Softmax benchmarks