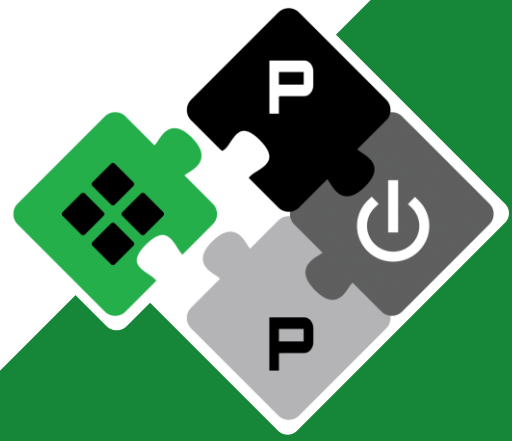
ETH zürich ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

# CachePool: Many-core cluster of customizable, lightweight scalar-vector PEs for irregular L2 data-plane workloads


Integrated Systems Laboratory (ETH Zürich)

**Zexin Fu, Diyou Shen** zexifu, dishen@iis.ee.ethz.ch

**Alessandro Vanelli-Coralli** avanelli@iis.ee.ethz.ch
**Luca Benini** lbenini@iis.ee.ethz.ch

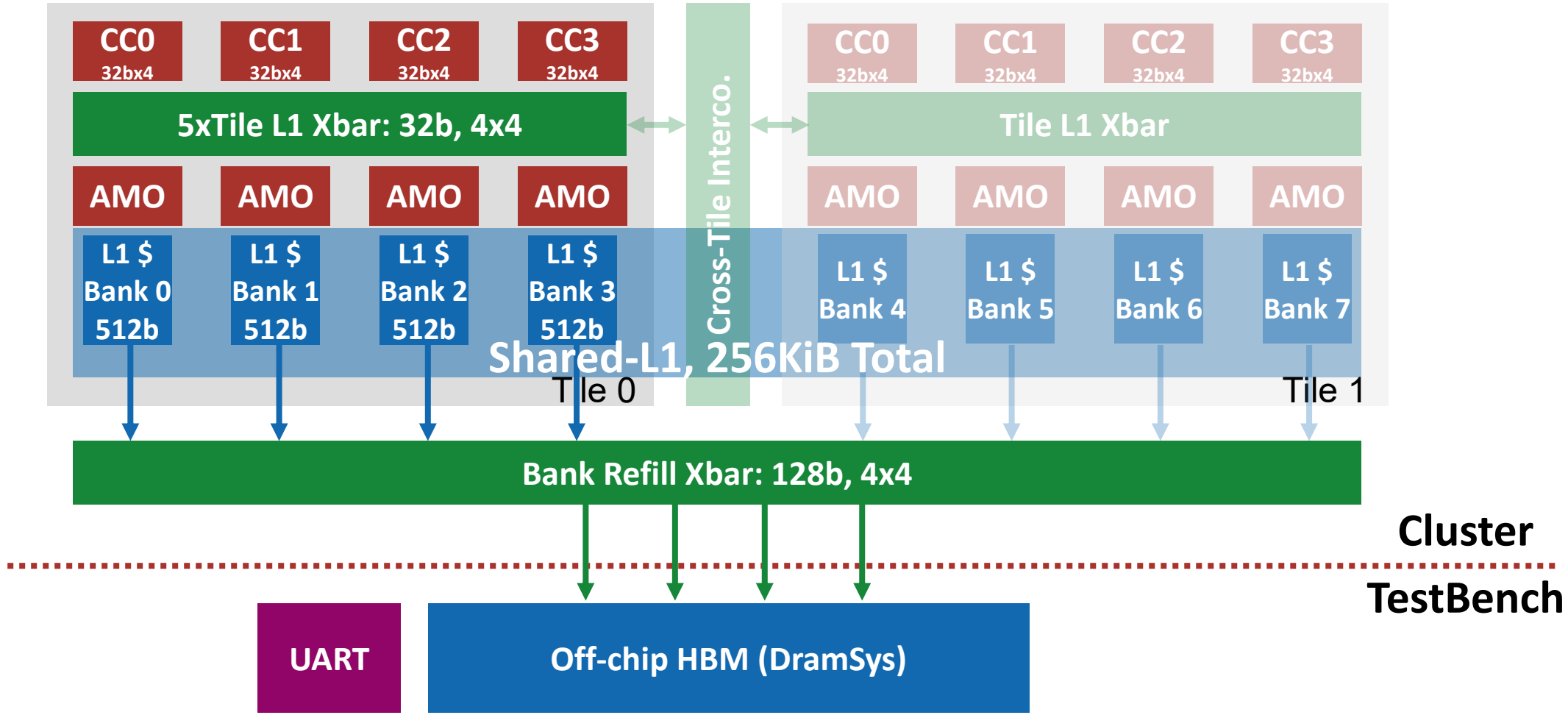
@pulp_platform

pulp-platform.org

youtube.com/pulp_platform

**PULP Platform**
Open Source Hardware, the way it should be!

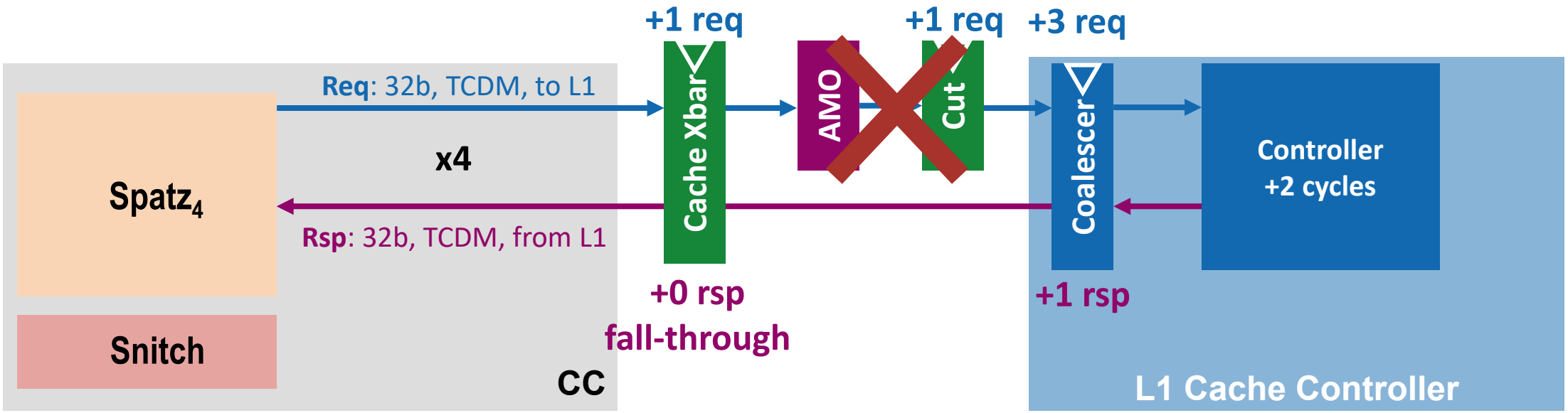
# Hardware Development



# Hardware Development

- **Optimize the hit-to-use latency for both Spatz and Snitch**
- **Draft a Roadmap to scale up => to be discussed with Luca in this week**
- **WIP: Add byte-write to InSitu Cache**

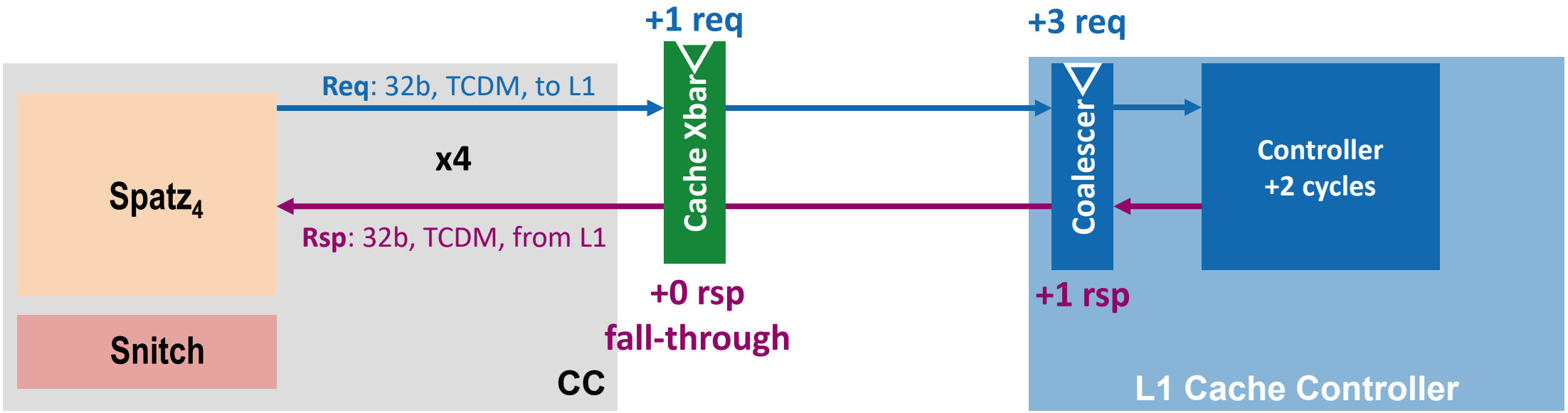
# Hit-to-Use Latency Optimization -- Spatz




**In total: 1 + 1 + 3+ 2 + 1 = 8 Cycles**

- **Interconnection adds 2 cycles on req: 1 for xbar critical path cut, 1 for atomic units**
- **Coalescer adds 1~3 cycles on request, 1 cycle on response**
  **=> Depends on cacheline width, needs optimization**
- **Controller hit takes 2 cycles**

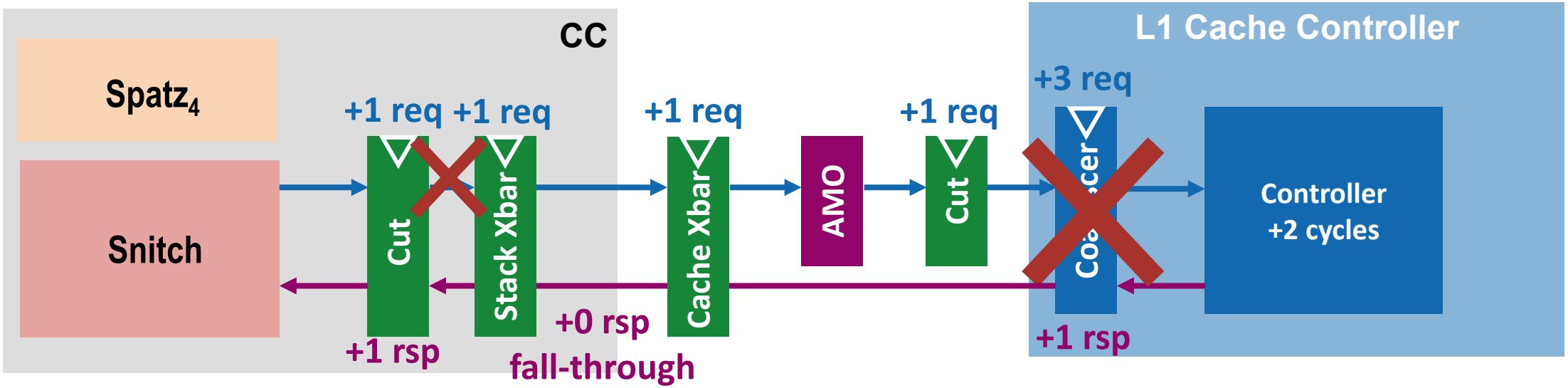
# Hit-to-Use Latency Optimization -- Spatz




**In total: 1 + 3+ 2 + 1 = 7 Cycles**

- **Remove AMO and CUT on Spatz path => No Vector AMO instruction**
- **Reduce 1 cycle**
- **Backend verified: timing close @1GHz, WW corner**

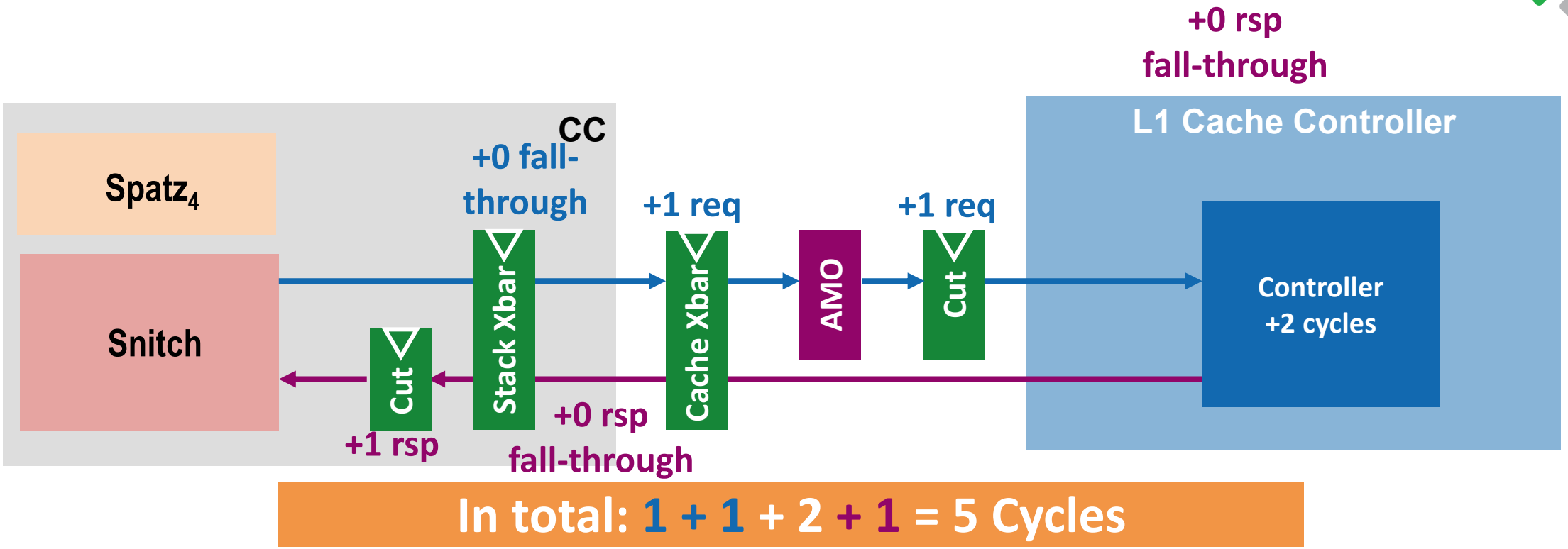
# Hit-to-Use Latency Optimization -- Snitch




**In total: 1 + 1 + 1 + 1 + 3 + 2 + 1 + 1 = 11 Cycles**

- **Unnecessary cuts on request path**
- **Coalescer not need for Snitch => no entry to combine with**

# Hit-to-Use Latency Optimization -- Snitch

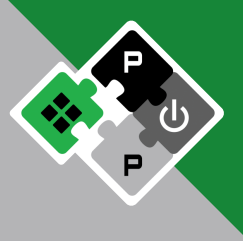



**In total: 1 + 1 + 2 + 1 = 5 Cycles**

- **Unnecessary cuts on request path**
- **Coalescer not need for Snitch => no entry to combine**
- **Backend verified: Ongoing, may need to adjust RSP cut position**

# Current Plan

- **Finish backend verification on hit-to-use latency optimization**
- **Confirm the calendar and scaling plan with Luca**
  - Will present in our next meeting
- **Add CI flow**
- **Finish byte-write support**



# Thank you!