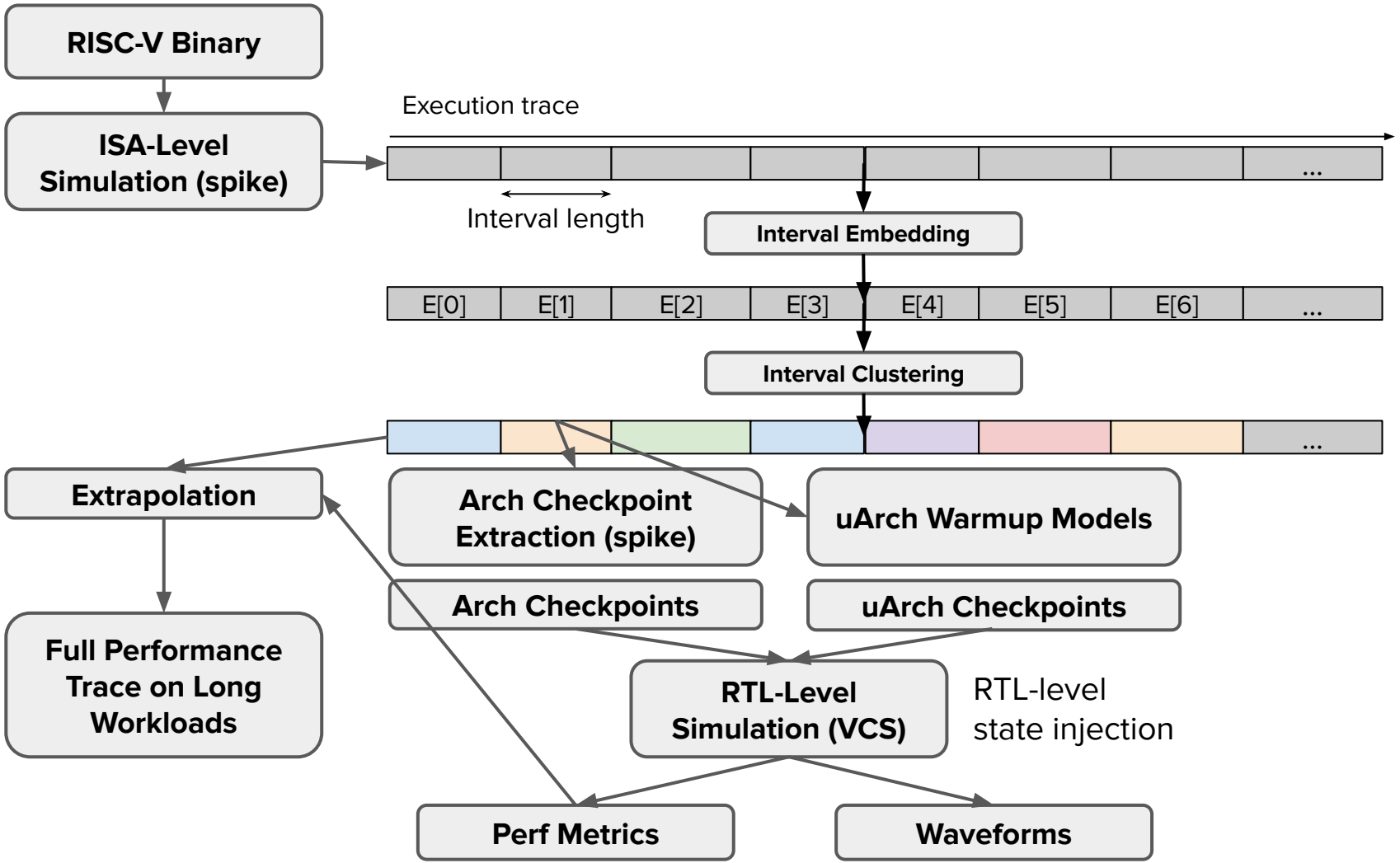
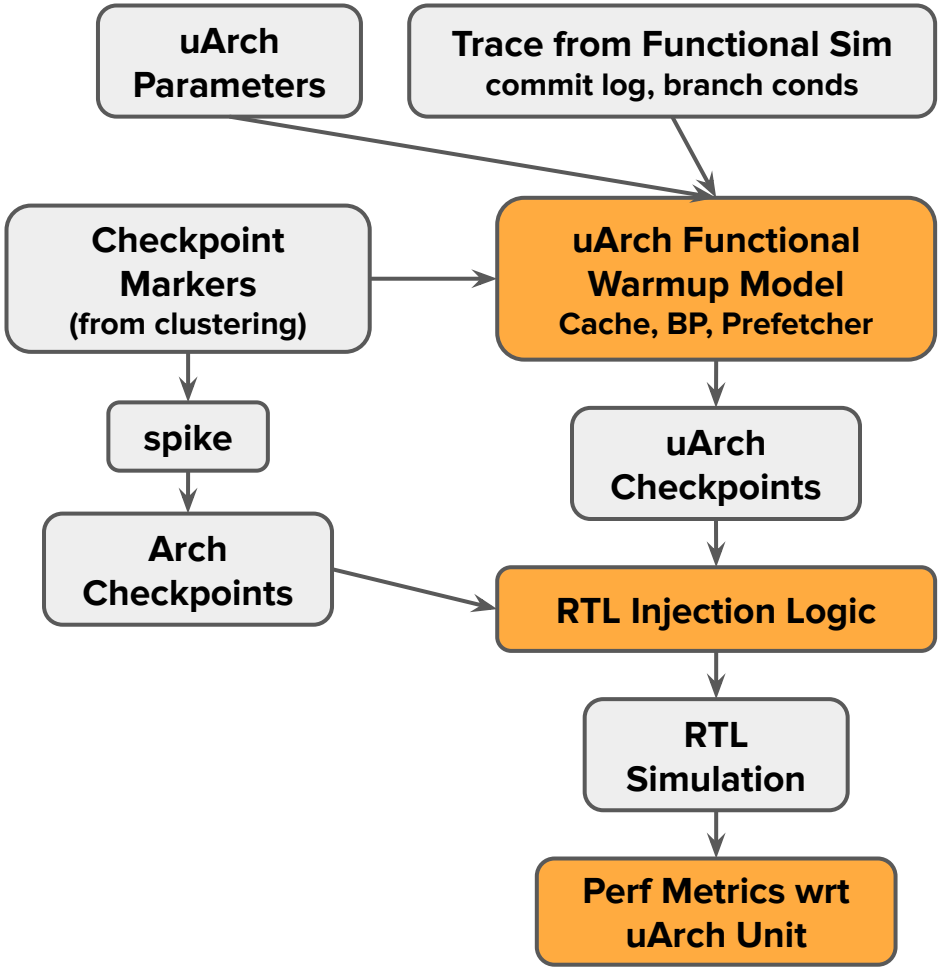
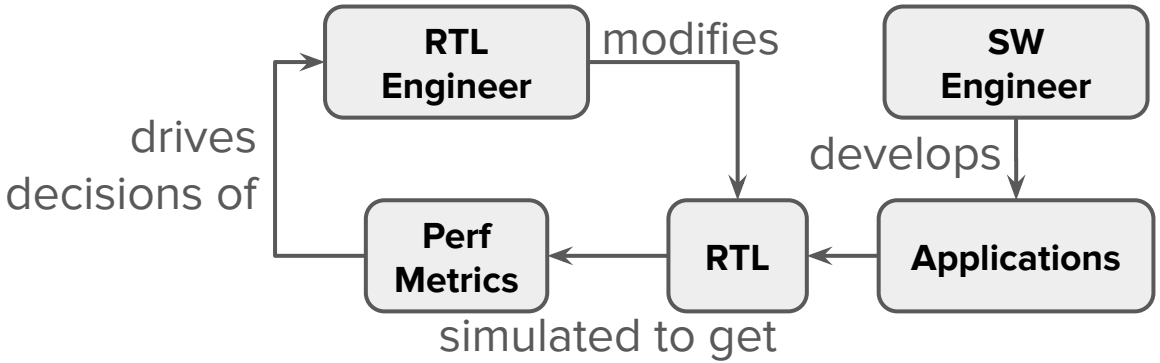
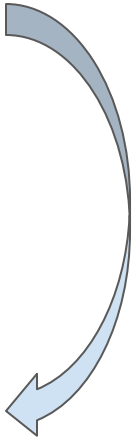

RISC-V Binary
ISA-Level Simulation (spike)
Execution trace
...
Interval length
Interval Embedding
E[0]
E[1]
E[2]
E[3]
E[4]
E[5]
E[6]
...
Interval Clustering
...
Extrapolation
Arch Checkpoint Extraction (spike)
uArch Warmup Models
Arch Checkpoints
uArch Checkpoints
Full Performance Trace on Long Workloads
RTL-Level Simulation (VCS)
RTL-level state injection
Perf Metrics
Waveforms


uArch Parameters
Trace from Functional Sim
commit log, branch conds
Checkpoint Markers
(from clustering)
uArch Functional Warmup Model
Cache, BP, Prefetcher
spike
uArch Checkpoints
Arch Checkpoints
RTL Injection Logic
RTL Simulation
Perf Metrics wrt uArch Unit


RTL Engineer
modifies
SW Engineer
drives decisions of
develops
Perf Metrics
RTL
Applications
simulated to get

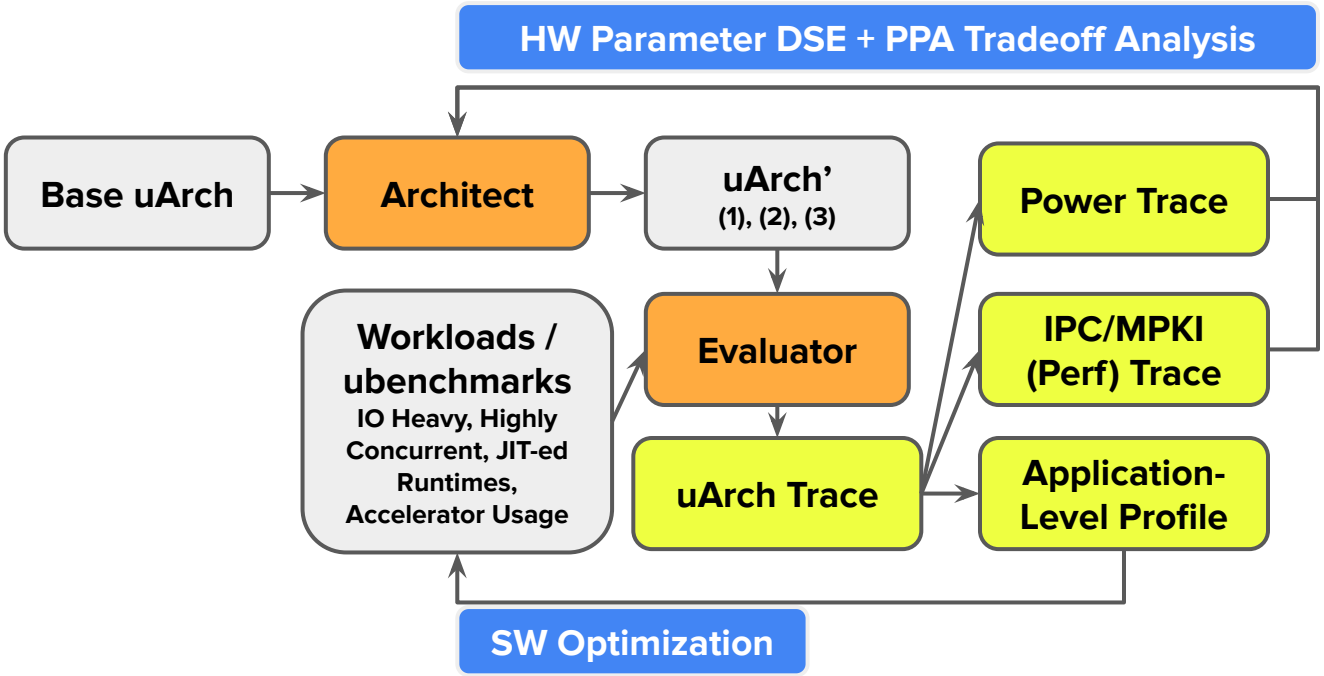
```
{address: 0x8000_0012, timestamp: 11, data: 0x0126, is_store: True,
writer_id: 1}

{address: 0x8000_0100, timestamp: 15, data: 0x0235, is_store: False,
writer_id: 1}
```

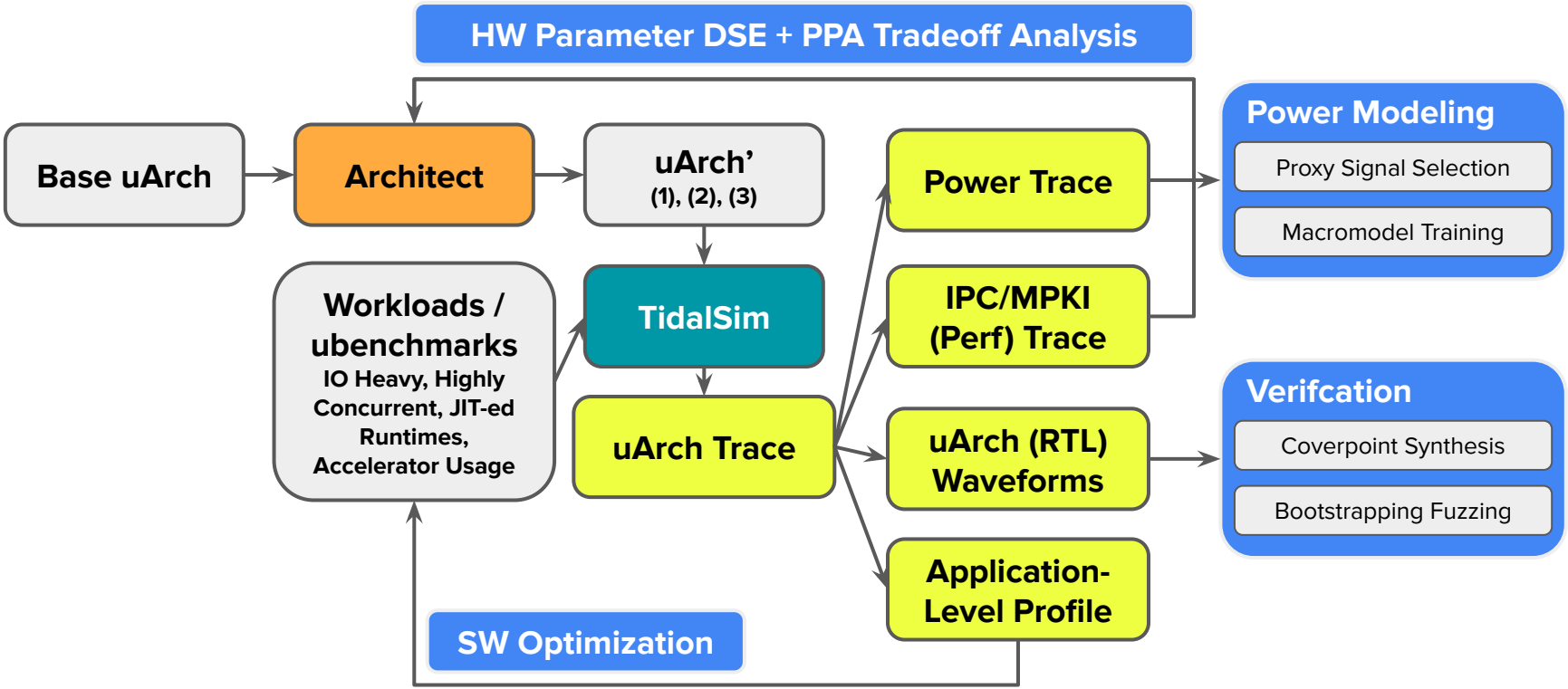


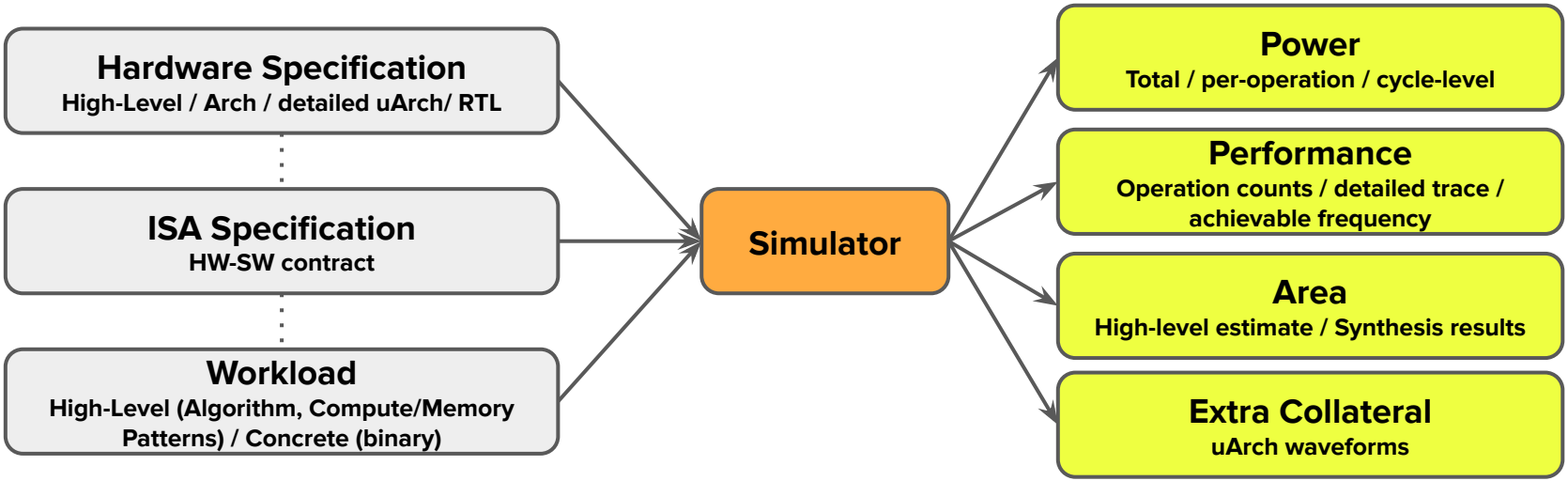
Address	Last Read Timestamp	Last Write Timestamp	Writer ID	Last Updated Value
0x8000_0010	10	~~5~~ 11	1	0x012600fc
0x8000_0100	15	7	1	0x00000235
0x8000_0020		12	1	0x00000140
⋮				

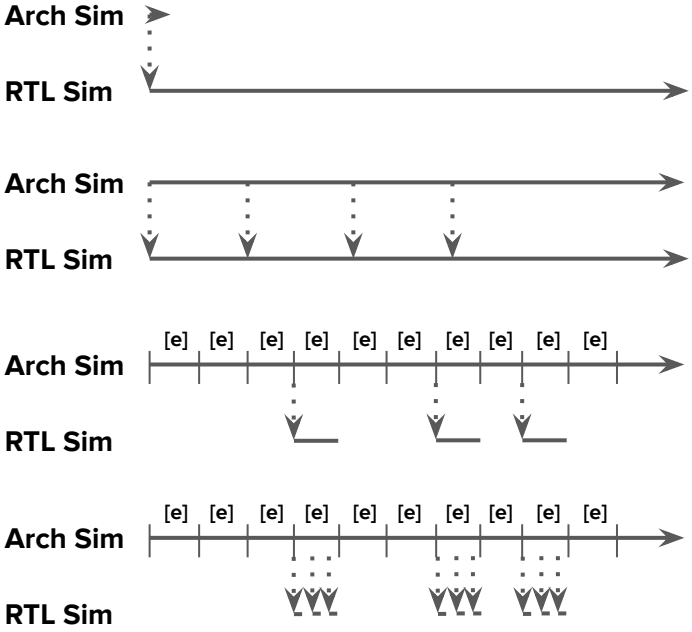
4B cache lines, 2-way set associative, 16 sets, write allocate, LRU eviction

Address	Last Read Timestamp	Last Write Timestamp	Writer ID	Last Updated Value
0x8000_0010	10	11	1	0x012600fc
0x8000_0100	15	7	1	0x00000235
0x8000_0020		12	1	0x00000140
⋮				
0x8000_0014	32	22	1	0x00000098


HW Parameter DSE + PPA Tradeoff Analysis
Base uArch
Architect
uArch'
(1), (2), (3)
Power Trace
Workloads / ubenchmarks
IO Heavy, Highly Concurrent, JIT-ed Runtimes, Accelerator Usage
Evaluator
IPC/MPKI (Perf) Trace
uArch Trace
Application-Level Profile
SW Optimization


HW Parameter DSE + PPA Tradeoff Analysis
Base uArch
Architect
uArch'
(1), (2), (3)
Workloads / ubenchmarks
IO Heavy, Highly Concurrent, JIT-ed Runtimes, Accelerator Usage
TidalSim
uArch Trace
Power Trace
IPC/MPKI (Perf) Trace
uArch (RTL) Waveforms
Application-Level Profile
SW Optimization
Power Modeling
Proxy Signal Selection
Macromodel Training
Verifcation
Coverpoint Synthesis
Bootstrapping Fuzzing


Hardware Specification
High-Level / Arch / detailed uArch/ RTL
ISA Specification
HW-SW contract
Workload
High-Level (Algorithm, Compute/Memory Patterns) / Concrete (binary)
Simulator
Power
Total / per-operation / cycle-level
Performance
Operation counts / detailed trace / achievable frequency
Area
High-level estimate / Synthesis results
Extra Collateral
uArch waveforms

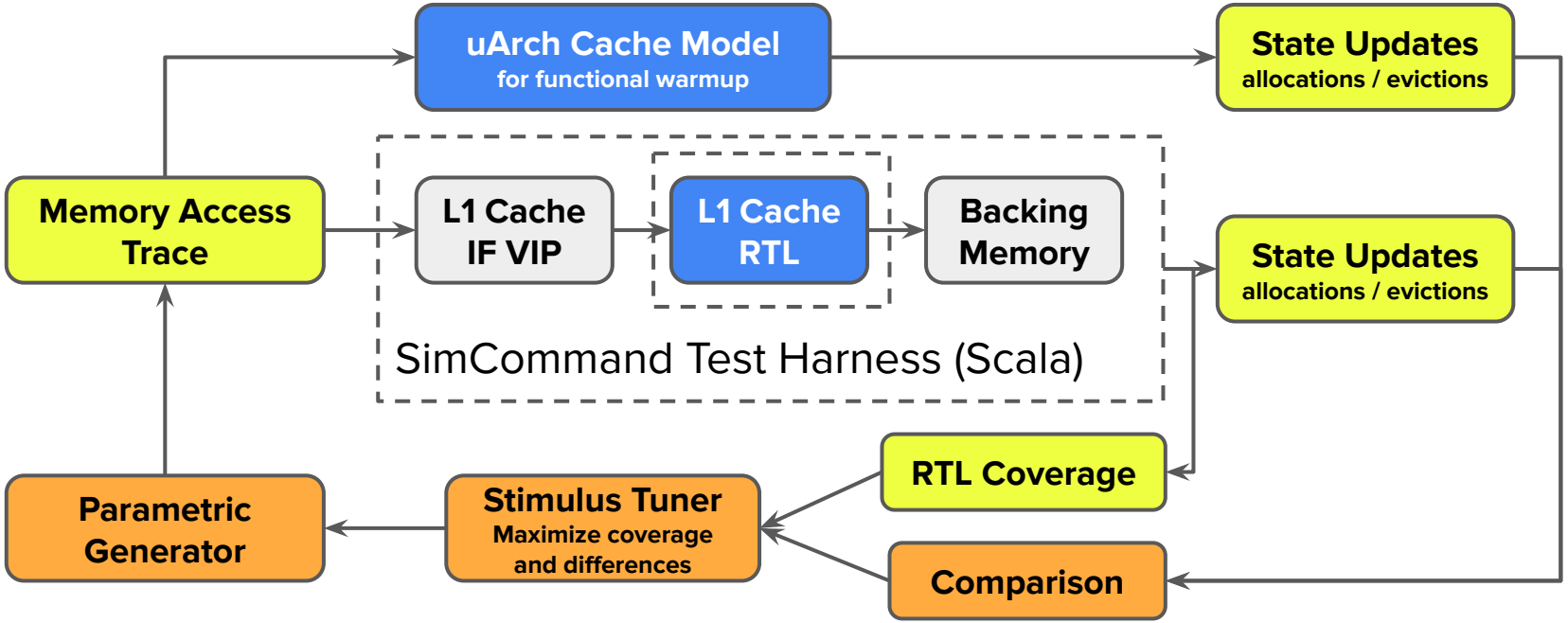


Full RTL Sim

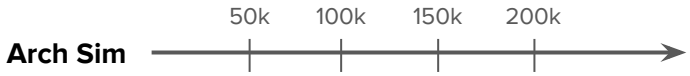
Parallel RTL Sim Dispatch with Functional Warmup

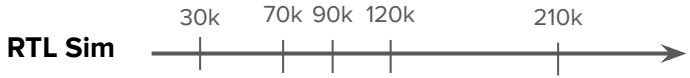
Representative Sampling

Representative Sampling with Random Subsampling

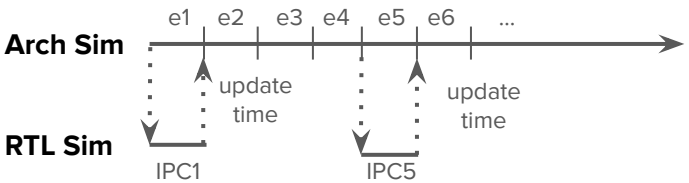

uArch Cache Model
for functional warmup
State Updates
allocations / evictions
Memory Access Trace
L1 Cache IF VIP
L1 Cache RTL
Backing Memory
State Updates
allocations / evictions
SimCommand Test Harness (Scala)
RTL Coverage
Parametric Generator
Stimulus Tuner
Maximize coverage and differences
Comparison

Timer interrupt points by dynamic instruction count

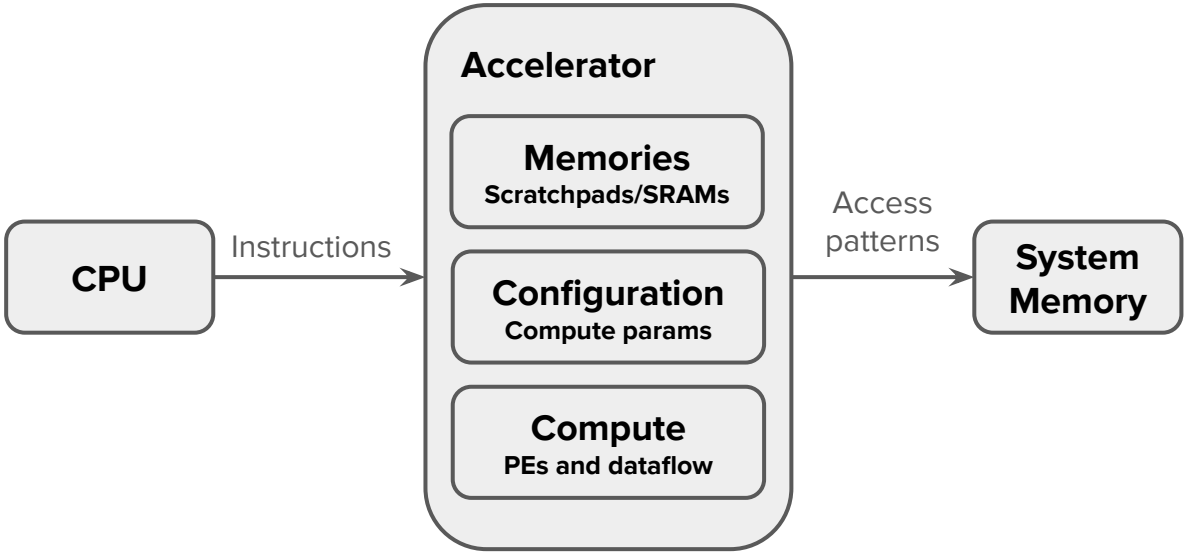




RTL and Arch sim increasingly drift out of sync!



Use feedback from RTL sim to arch sim to estimate time advancement via embedding similarity


Accelerator
Memories
Scratchpads/SRAMs
Configuration
Compute params
Compute
PEs and dataflow
CPU
Instructions
Access
patterns
System
Memory