PCIeHLS

Malte Vesper, Dirk Koch and Khoa Pham

High level synthesis – half a solution

- Easy generation of kernels from popular languages
- Good results require tuning with knowledge about FPGA architecture
- No infrastructure for kernel

C++

C

OpenCL

C

Vendor kernel integration

- For selected boards only:
  => Popular academic board VC709 missing
- Vendor partial flow

Partial flow

Ours

✘ Potentially calls for minor manual adjustments on the static system

✓ Relocation of modules

✓ Combining partial regions

✓ Synthesis largely independent of static system

✓ Synthesis of partial and static with different tool versions

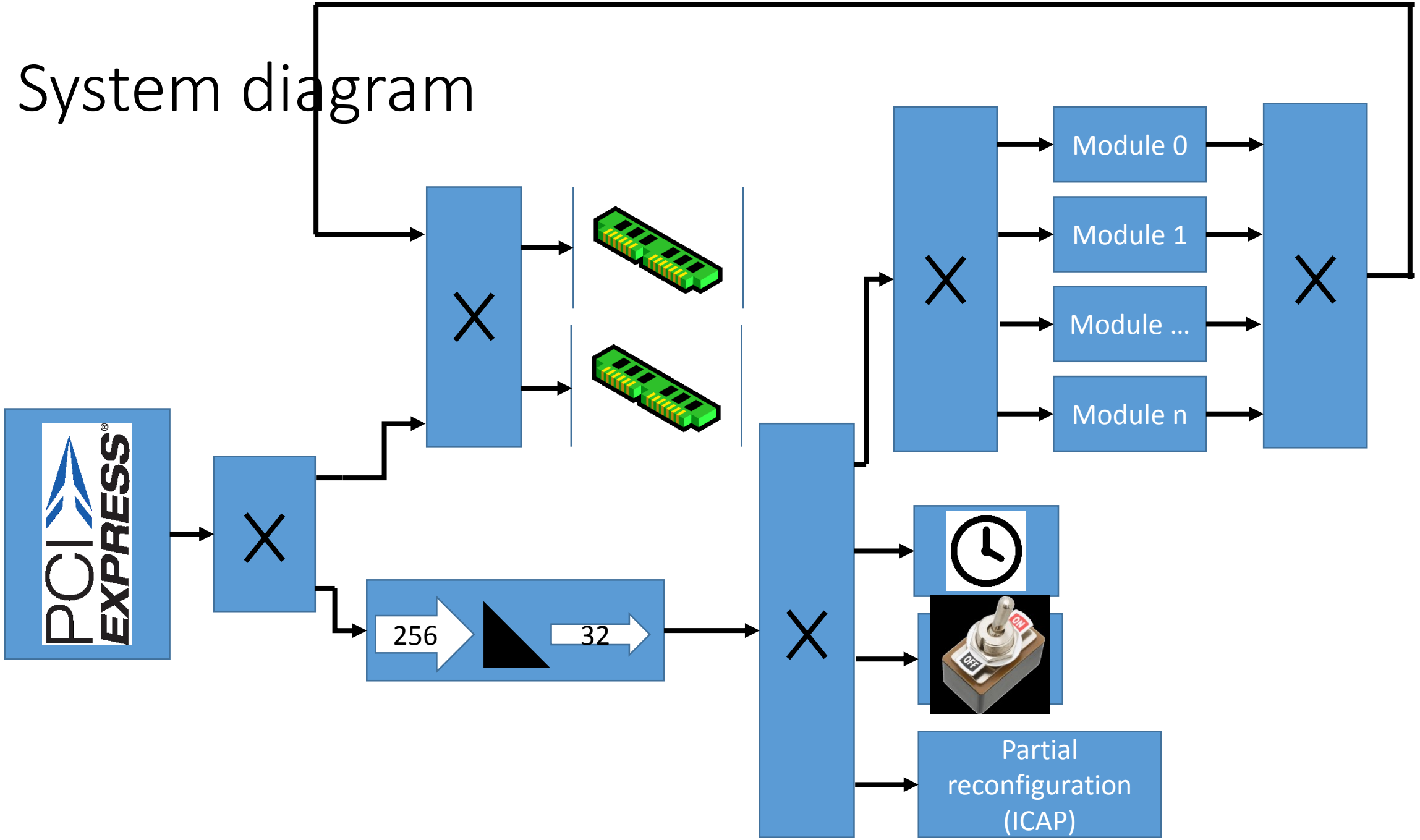
Xilinx

✓ Commercial stability

Things you don’t want to know

- ICAP
- PCIe
- Memory controller
- Decoupling
- Clock domain crossing
- Timing closure

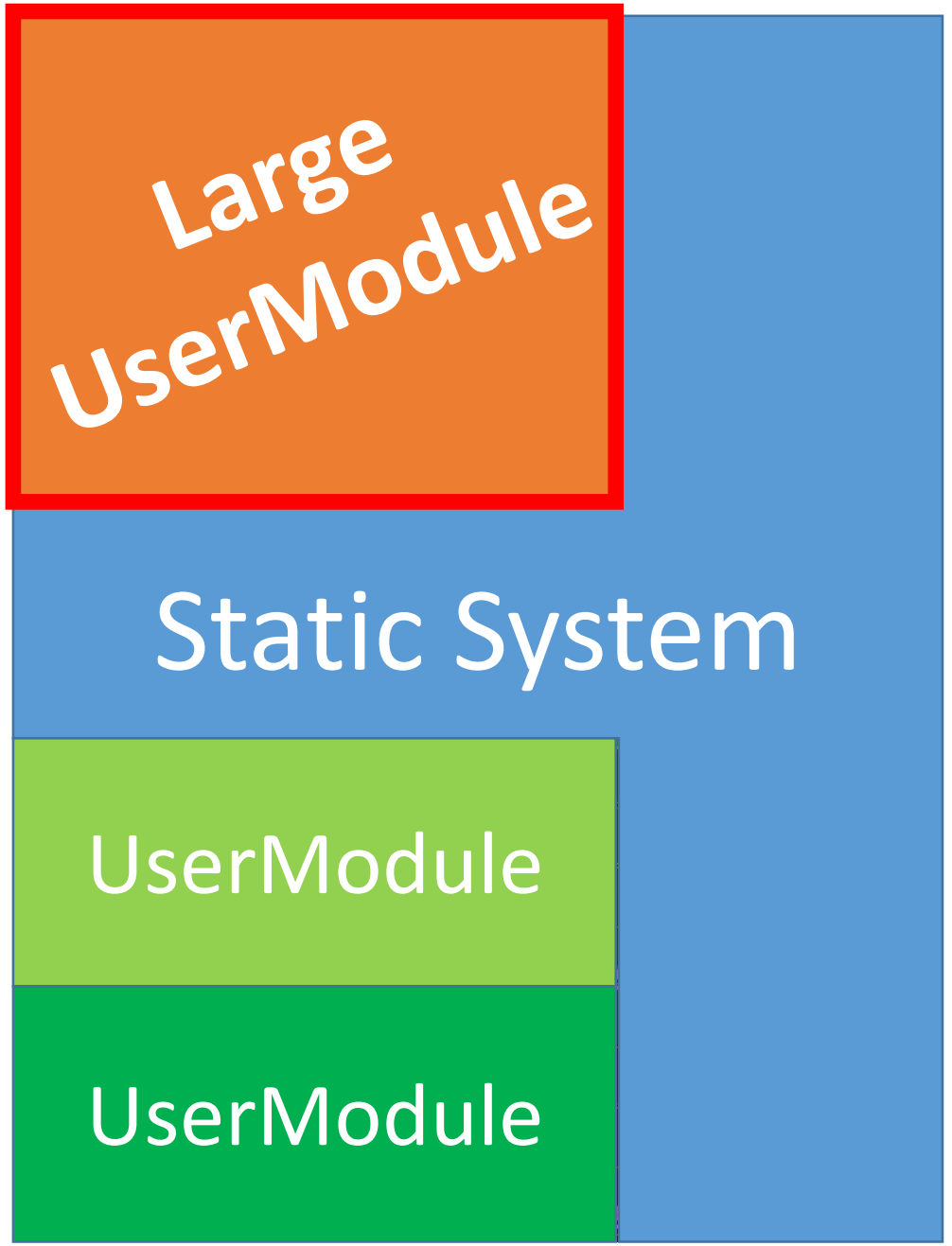
System diagram



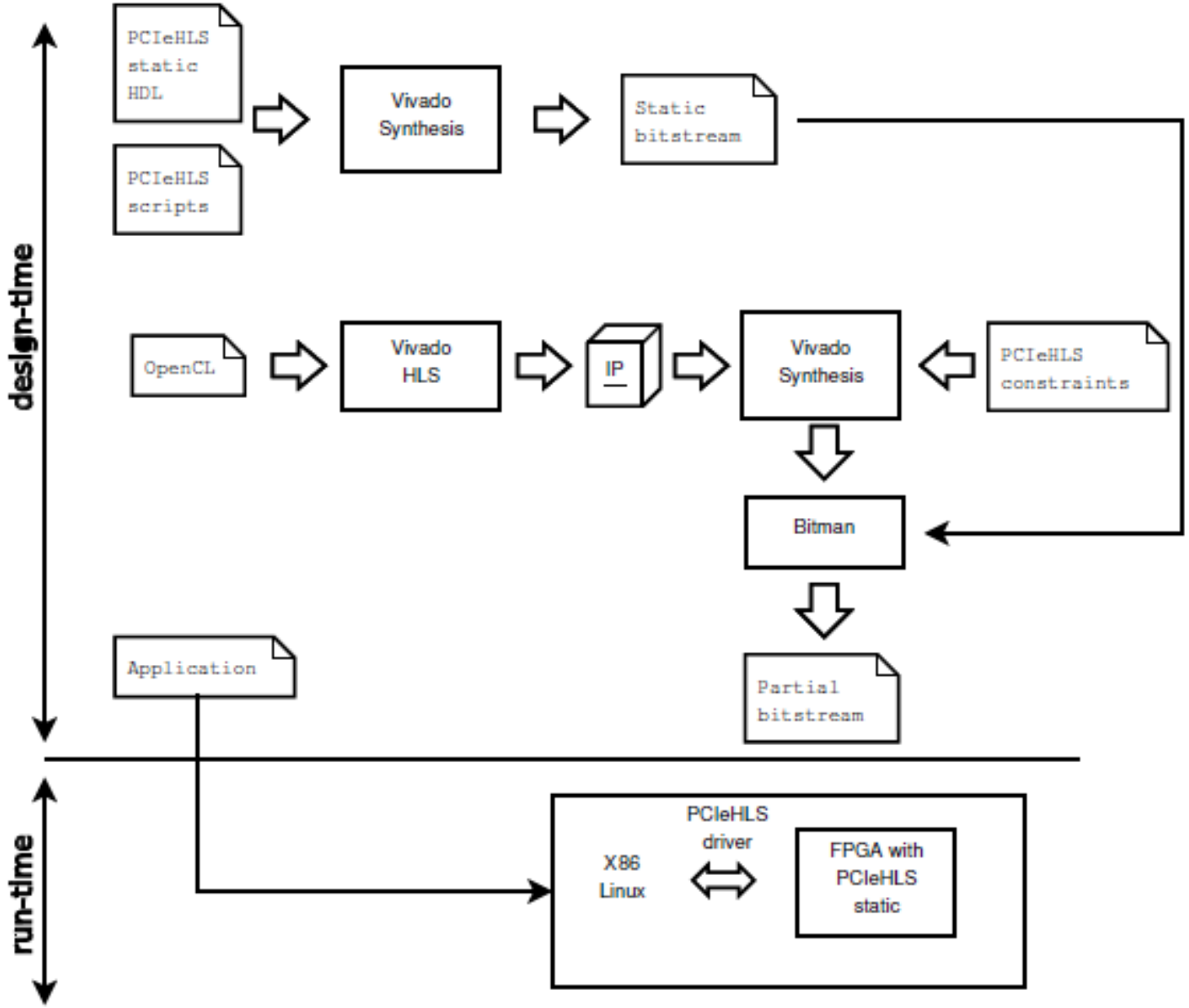
Floorplan

- Up to 4 user modules
- Each user module ≈13.5% Slices
- Static system ≈46.0% Slices

- Adjacent user module areas can be combined



Flow

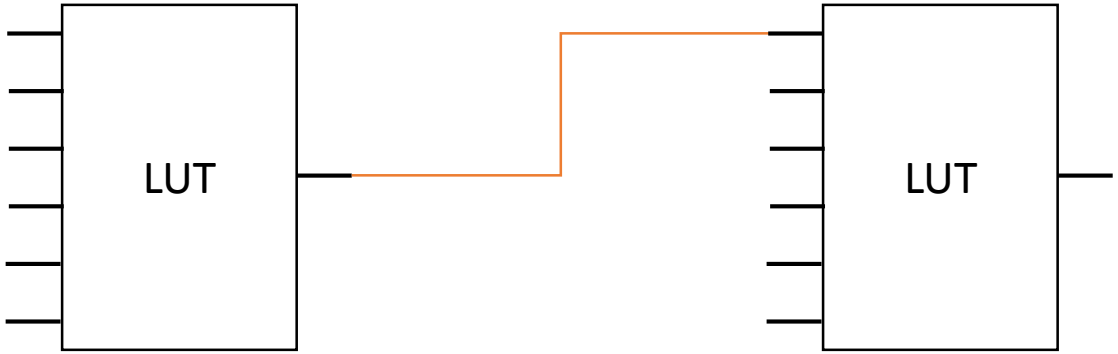

design-time
PCIeHLS static HDL
PCIeHLS scripts
Vivado Synthesis
Static bitstream
OpenCL
Vivado HLS
IP
Vivado Synthesis
PCIeHLS constraints
Bitman
Partial bitstream
Application
run-time
PCIeHLS driver
X86 Linux
FPGA with PCIeHLS static

Steps of our flow

- Bus macro
- Clock constraining
- Block:
  - Fabric differences
  - Sites used by static system
  - Pips used by static system
- Timing constraints
- Cut out bitstream with Bitman

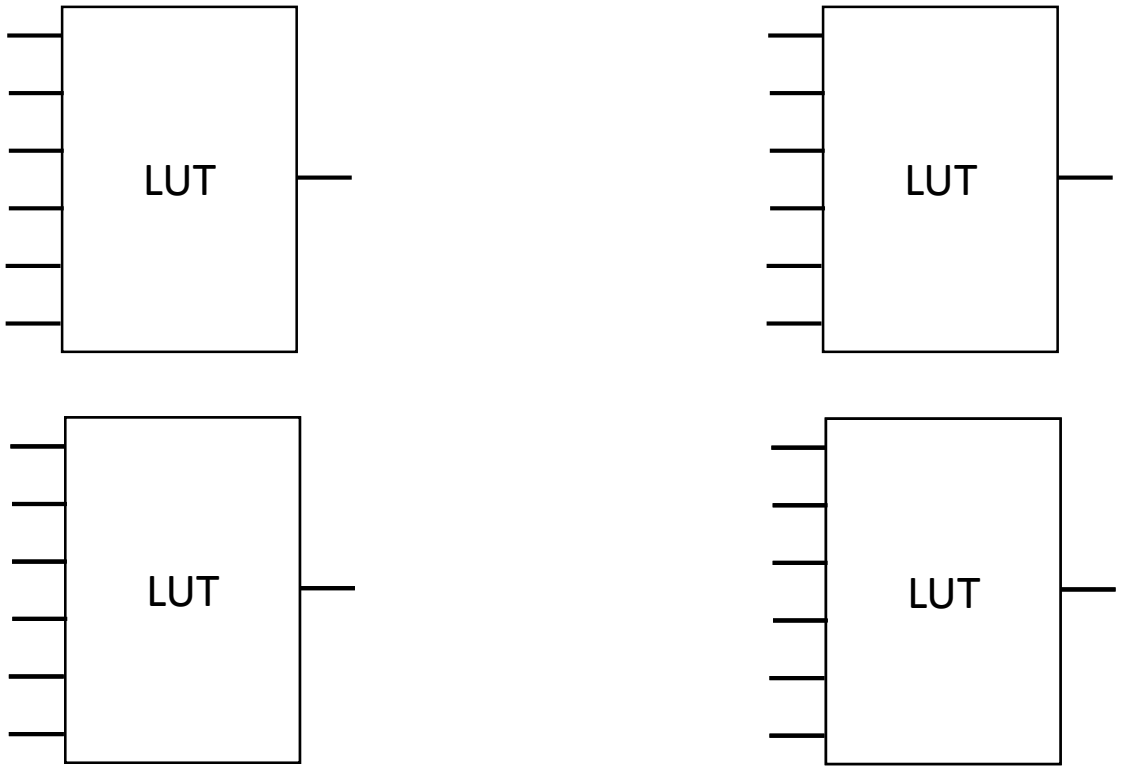
Bus Macro

- LUT – wire – LUT



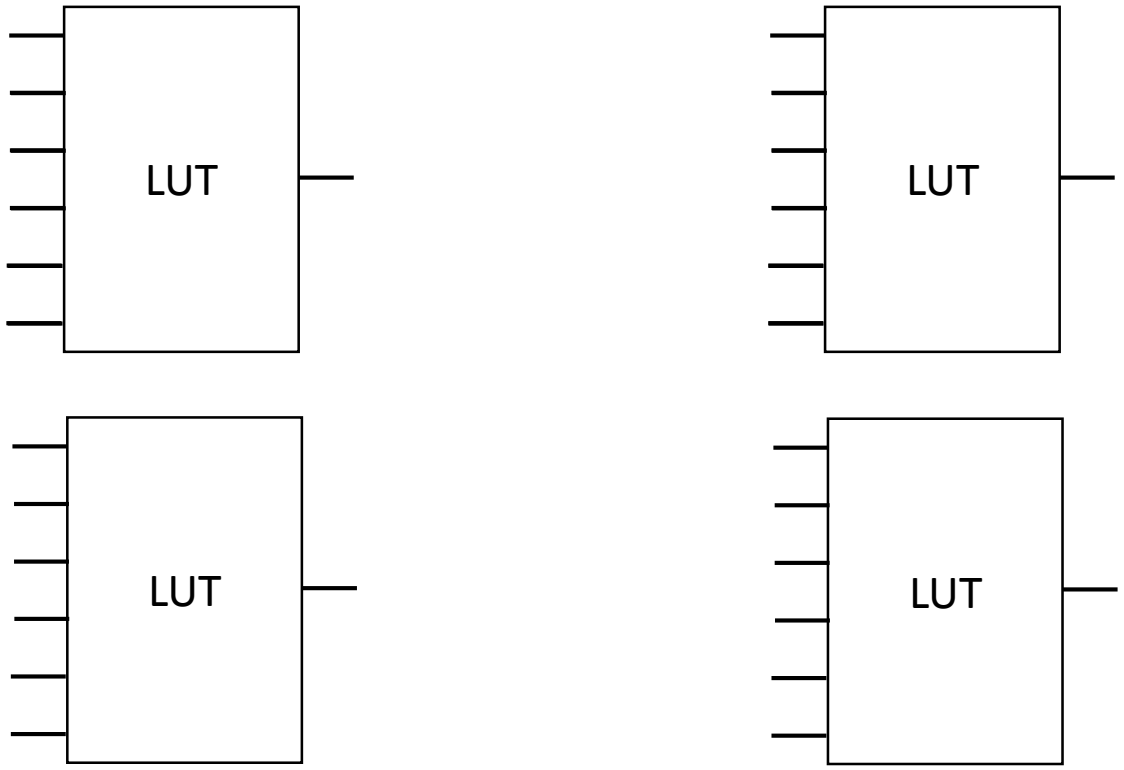
Bus Macro

- LUT – wire – LUT
- Constrained:
  - LOC/BEL



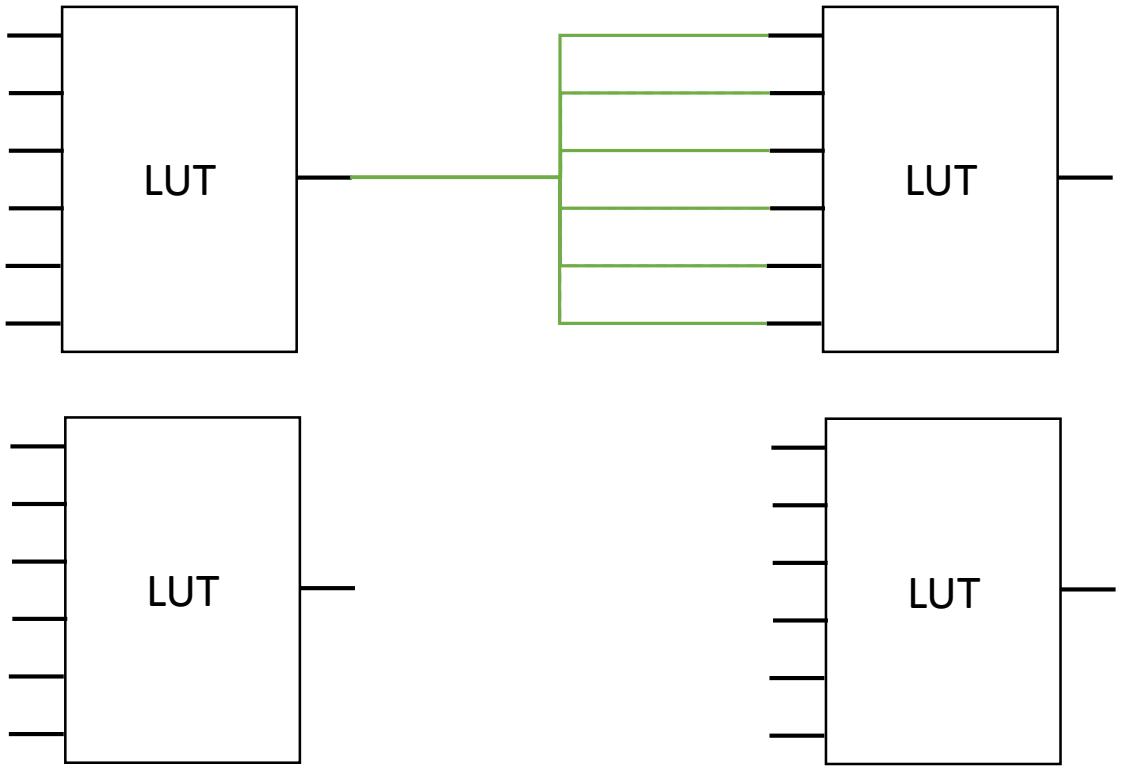
Bus Macro

- LUT – wire – LUT
- Constrained:
  - LOC/BEL
  - LOCK_PINS



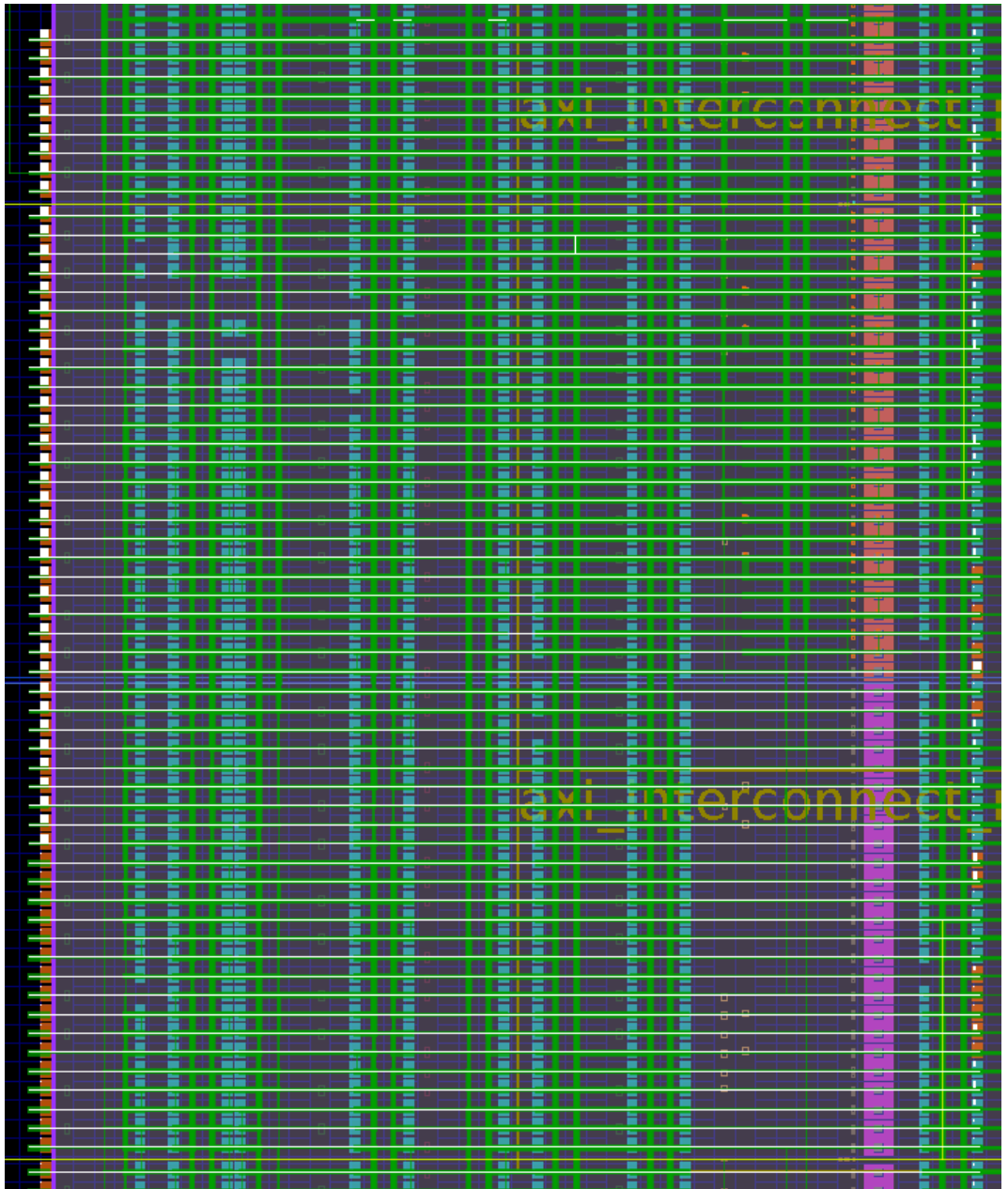
Bus Macro

- LUT – wire – LUT
- Constrained:
  - LOC/BEL
  - LOCK_PINS
  - FIXED_ROUTE



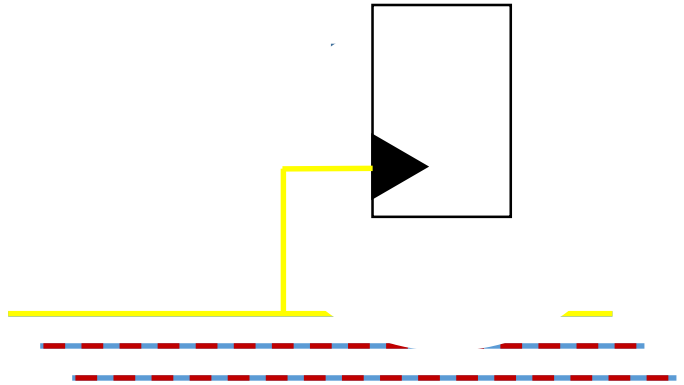
Bus Macro

- LUT – wire – LUT
- Constrained:
  - LOC/BEL
  - LOCK_PINS
  - FIXED_ROUTE



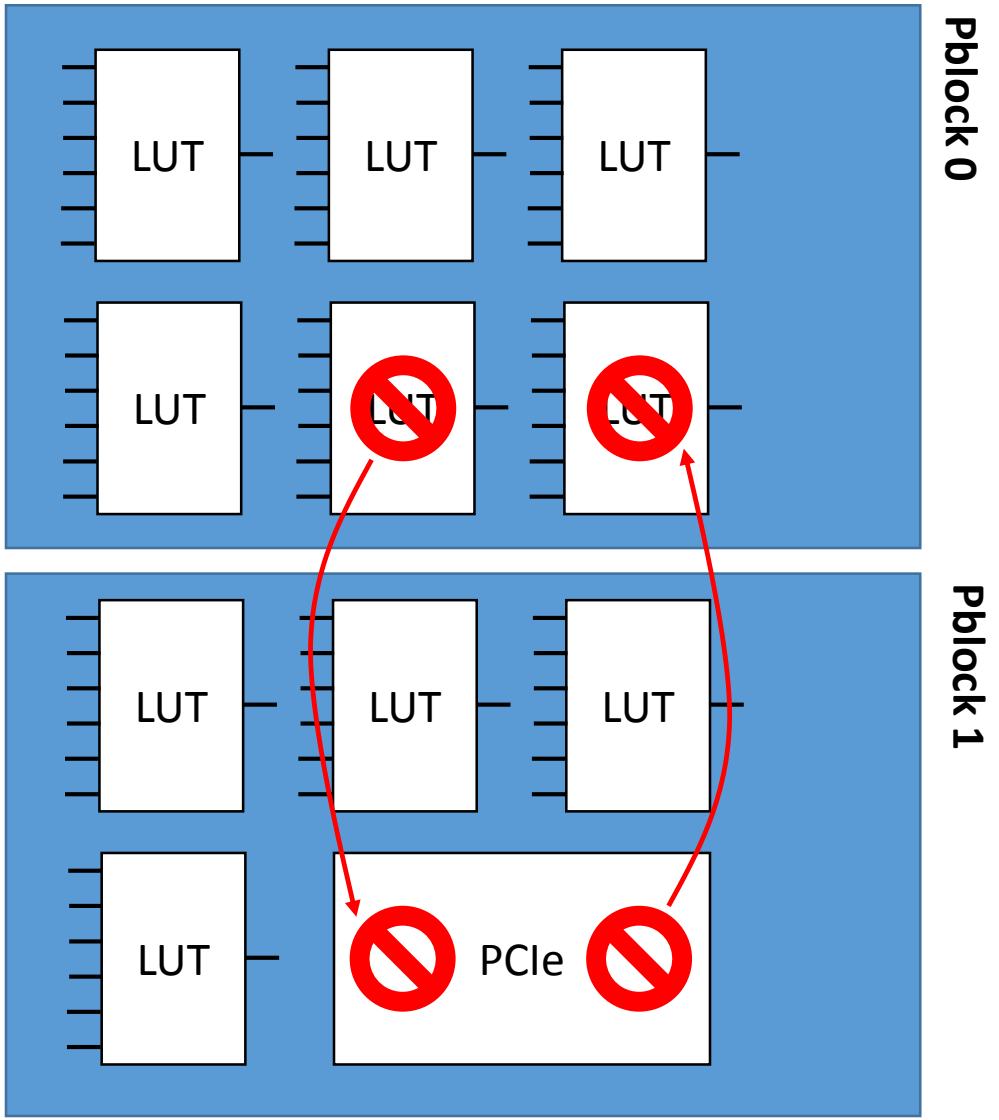
Clock constraining

- Ensure clock is driven
- Block other h-wires
- Issues: timing differences on relocation, positive and negative skew



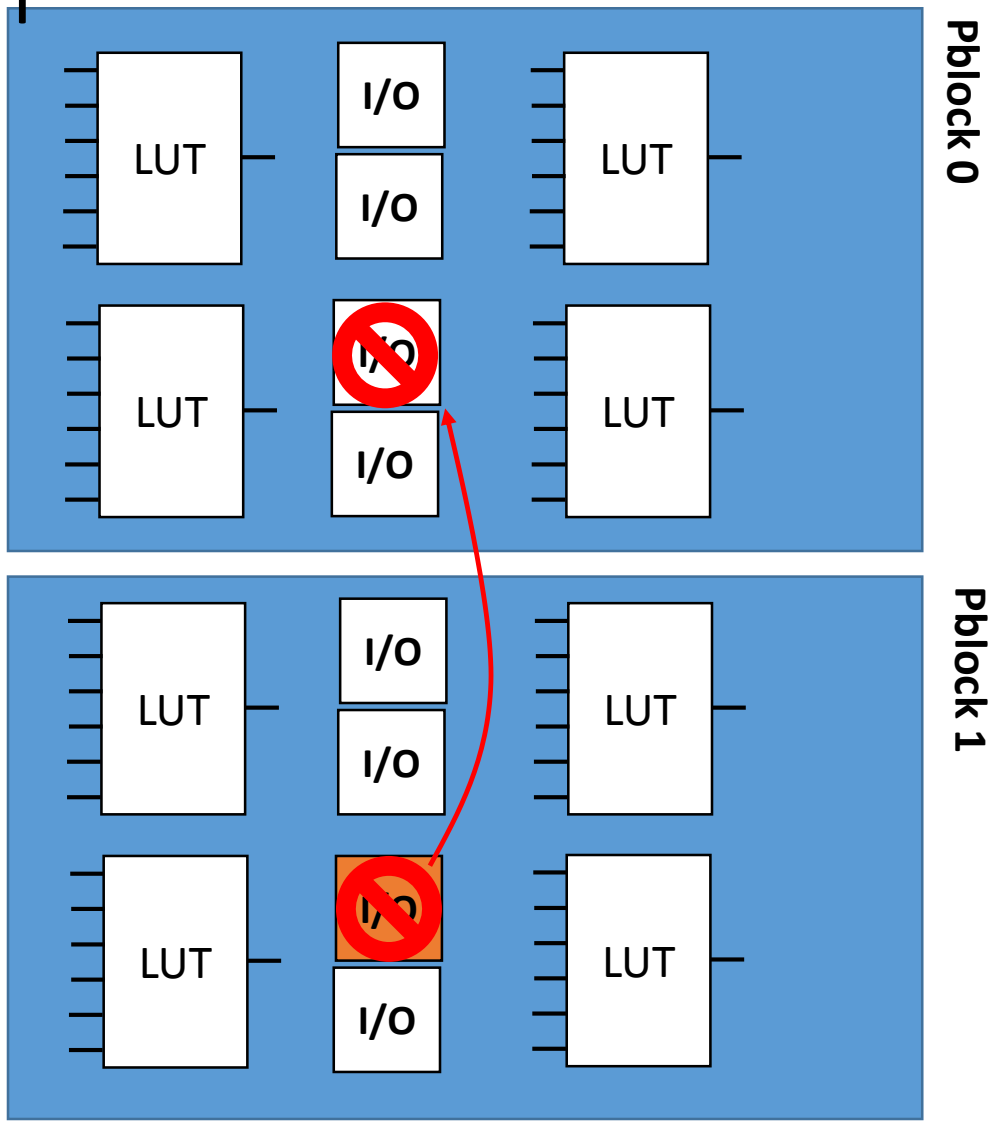
Fabric differences

- Special cells disturb regularity of fabric (i.e. PCIe, ICAP, …)
- Simply block differences



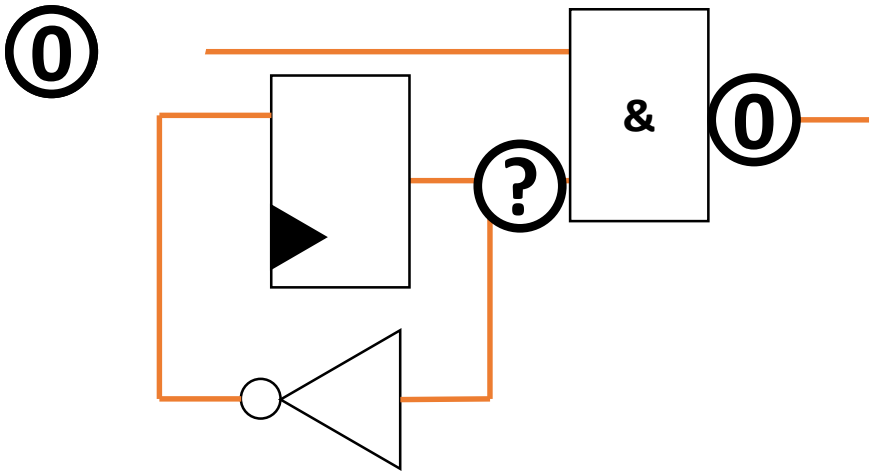
Sites used by static system

- Block
- I/O does not actually matter, not reconfigured



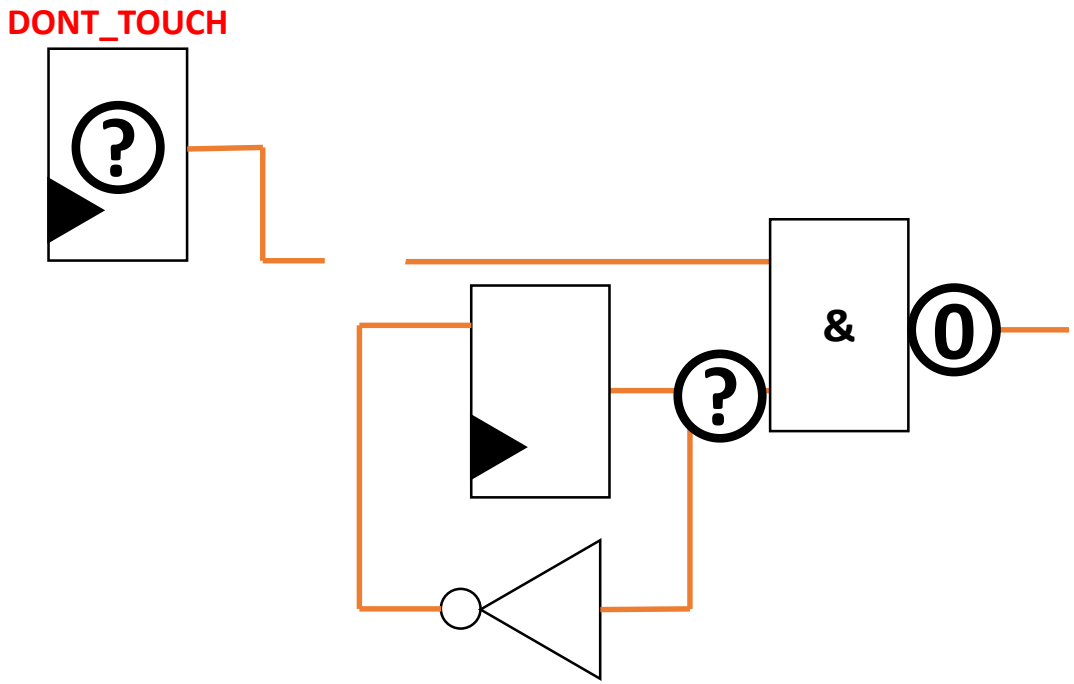
Optimization prevention

- Floating wires tied off to 0
- Optimization might remove logic

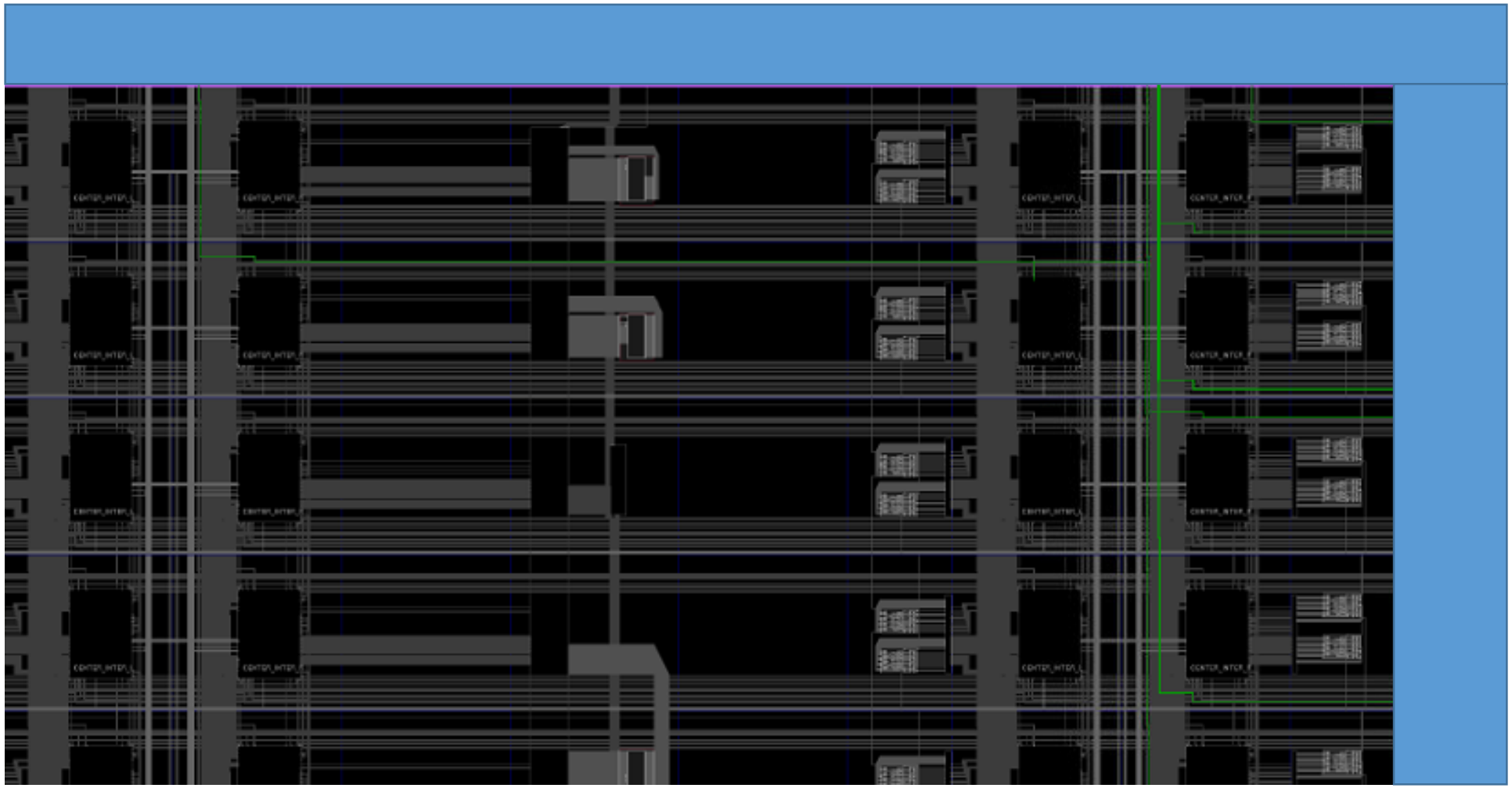


Optimization prevention

- Floating wires tied off to 0
- Optimization might remove logic
- Flop marked as DONT_TOUCH prevents logic optimization
- Works for signals into the partial region as well

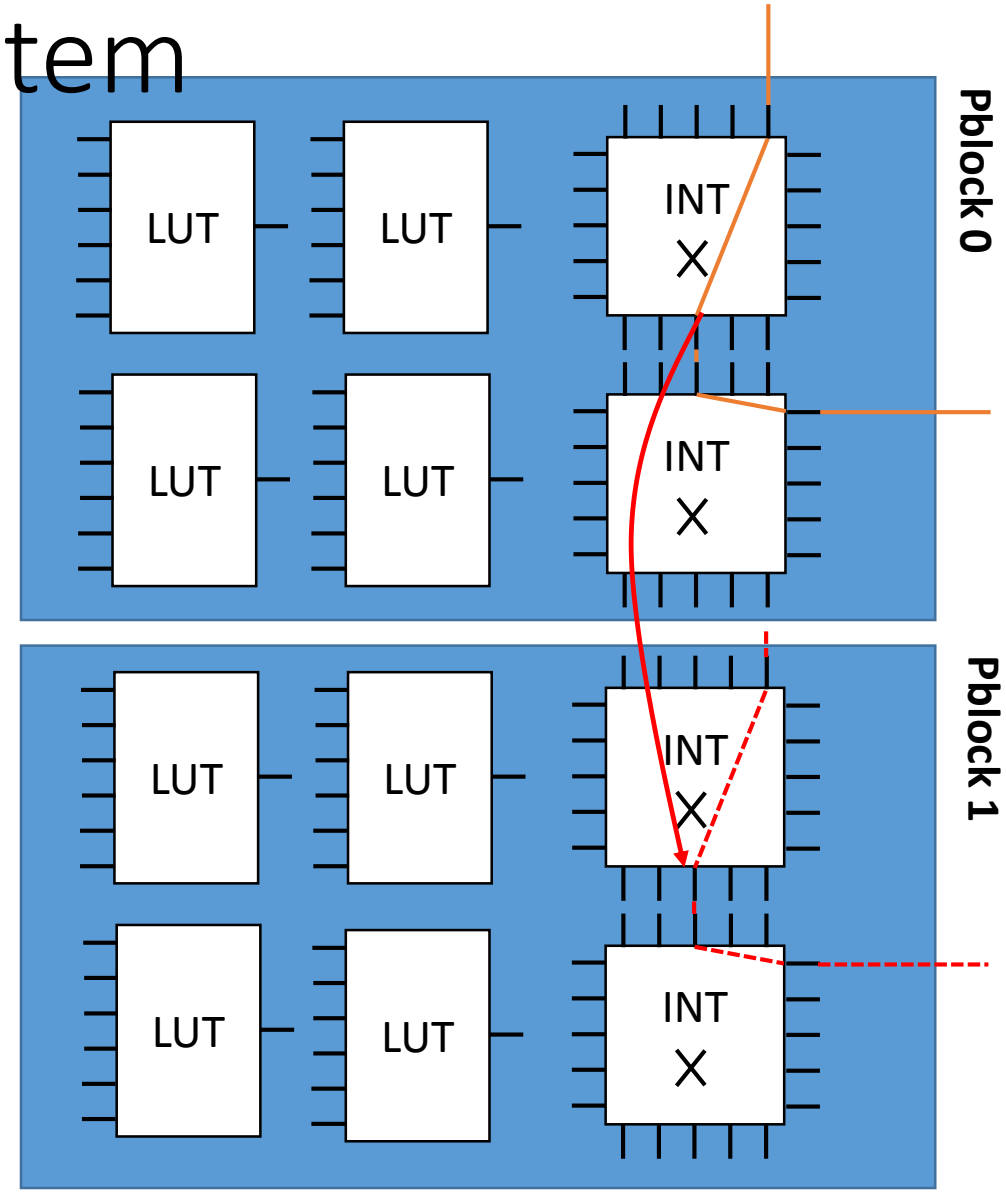


Routing used by static system



Routing used by static system

- Route a blocker from outside the PR through the wires (pips)

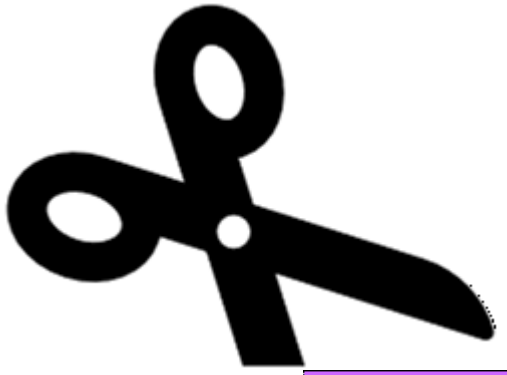


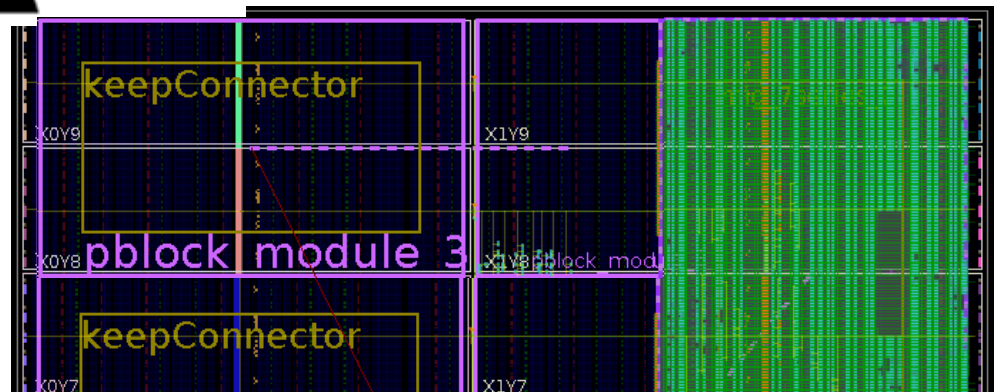
Timing constraints

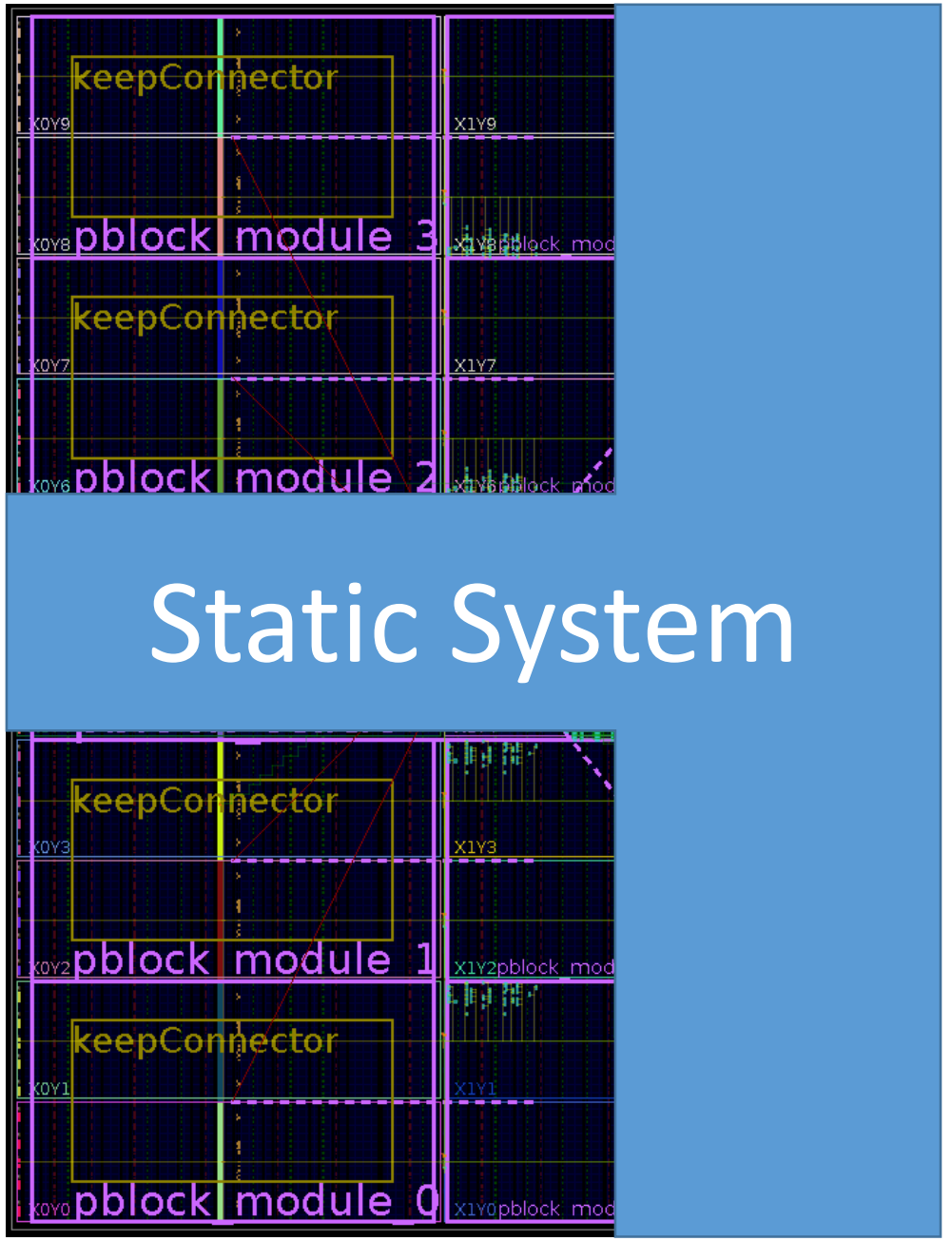
- Extract timing to Bus macro in static system
- Calculate slowest as WORST
- Constrain path of partial module to bus macro to period-worst

Bitman cutting

- extract partial bitstreams
- Relocate bitstreams for modules







Summary

- Build modules:
  - Once, use in multiple locations
  - Independent of static system
- Infrastructure provided:
  - ICAP partial reconfiguration
  - PCIe link to host
  - MMCM to adjust clock for partial modules
  - Memory

Thank you for your attention

Questions