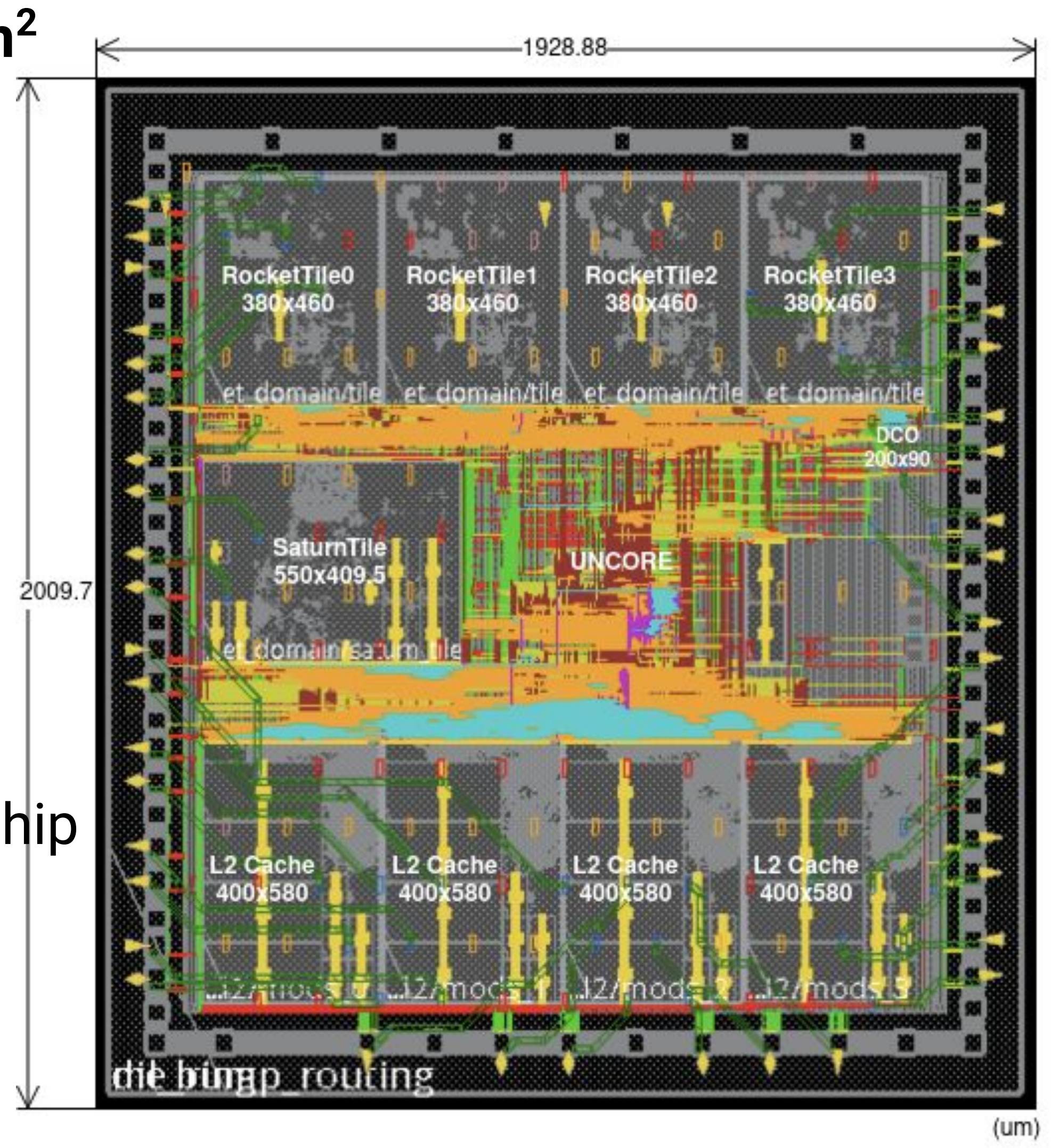
Intel 16 Tapeout - EE194/290C Spring 2022

Animesh Agrawal, Raghav Gupta, Roger Hsiao, Franklin Huang, Reza Sajadiany, Ella Schwarz, Jennifer Zhou

Bearly ML: SoC for Machine Learning

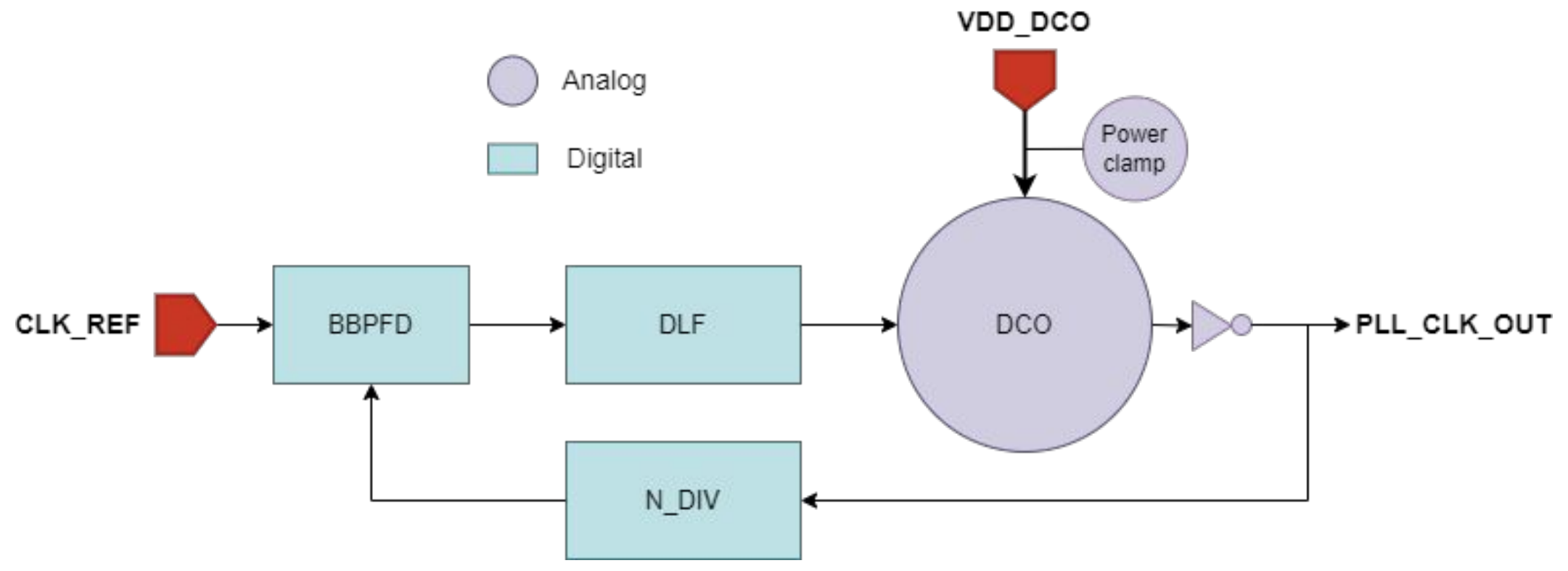
Chip Overview

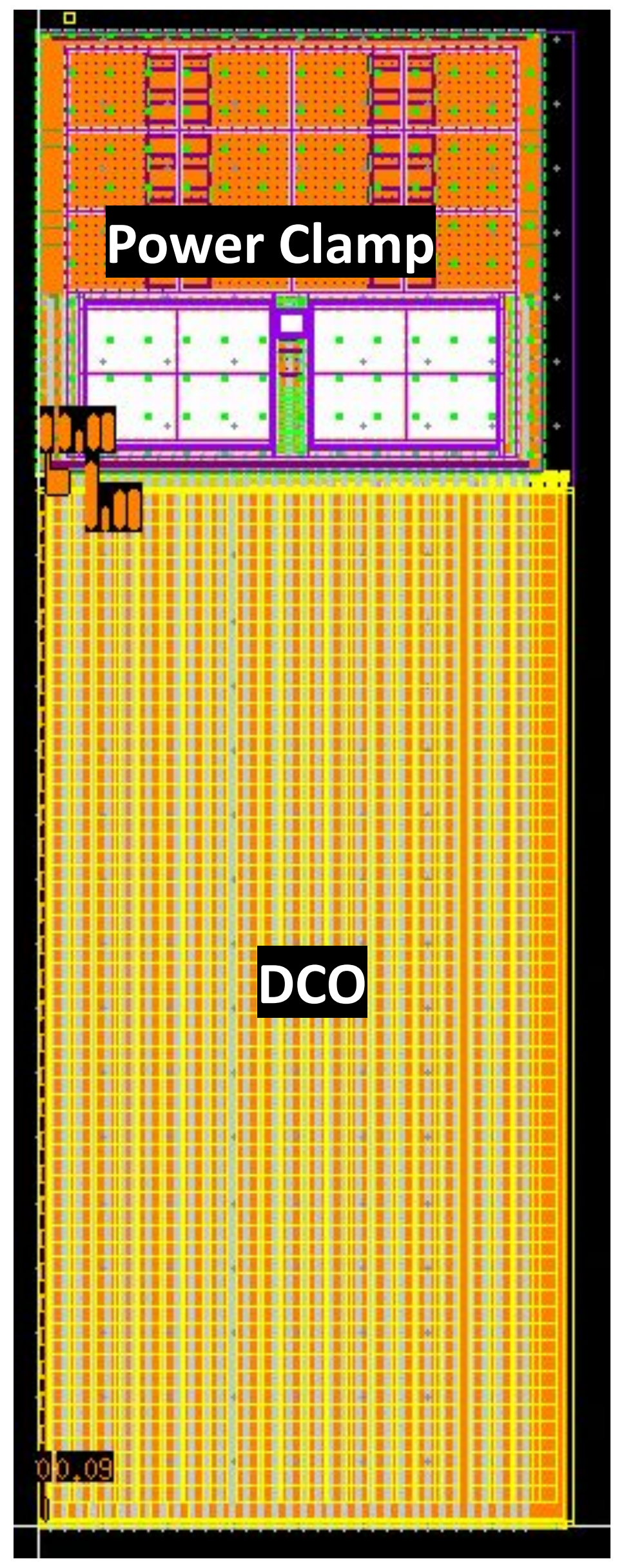
- RISC-V SoC for machine learning (ML) applications with ML/DSP accelerators in **Intel 16** technology
- Custom RTL, integrated into Chipyard template
- Area: **2009.7 x 1928.88 um^2**
- Core density: **72.46%**
- 1 general-purpose tile
- 4 specialized tiles
- Peripherals
  - SPI
  - I2C
  - UART
  - 4 GPIO pins
- 512 KB L2 cache
- Constellation network-on-chip
- PLL clock generation



Clocking

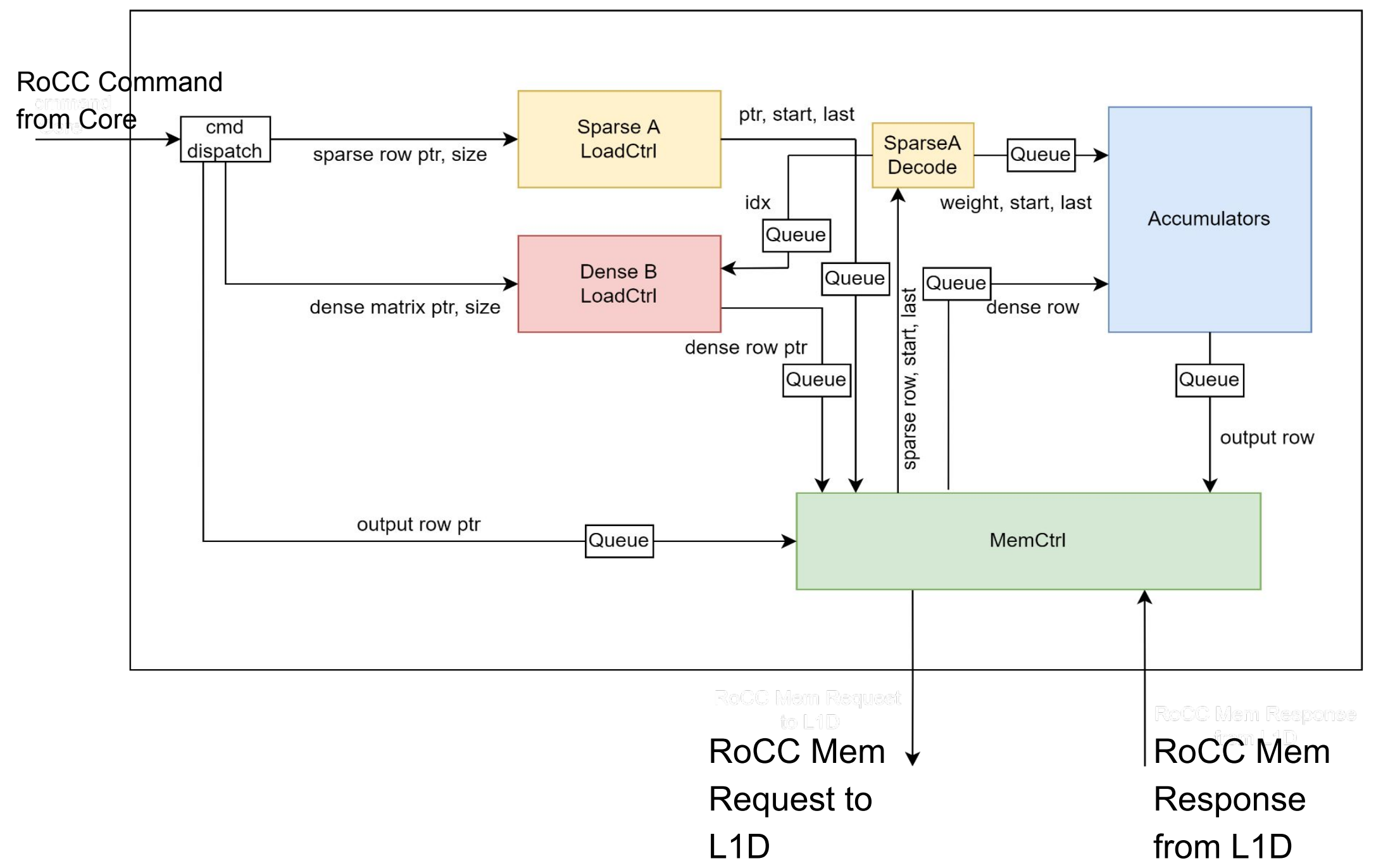
- Off-chip slow clock (< 100 MHz)
- On-chip fast clock generation with PLL (100 MHz - 1.2 GHz)
- Chip clock selected through MMIO
- Clock pinout to debug system clock
- Frequency of tile can be individually specified
- DCO generated using the Berkeley Analog Generator (BAG) designed by Sean Huang
- Digital logic is imported as black box verilog modules into Chipyard and placed close to the DCO to reduce parasitics

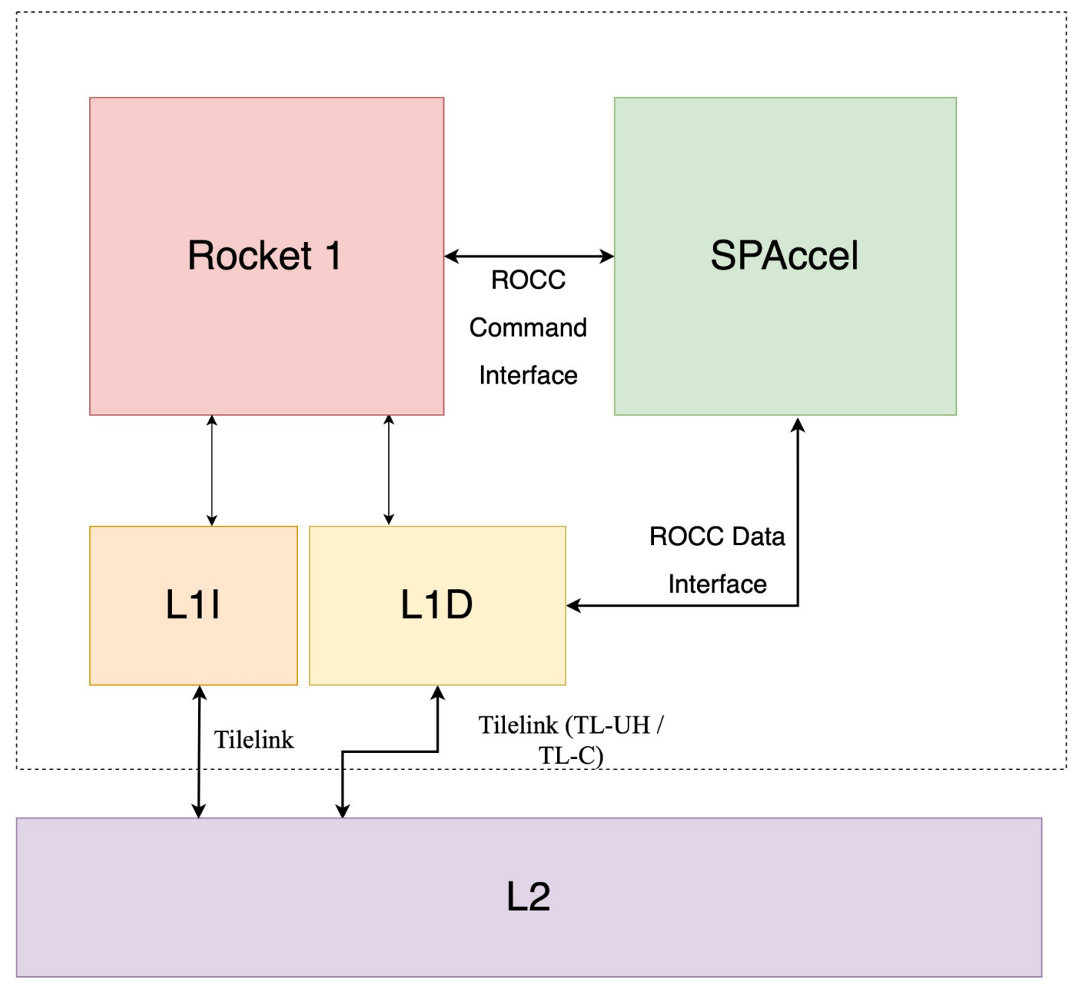




Specialized Tile

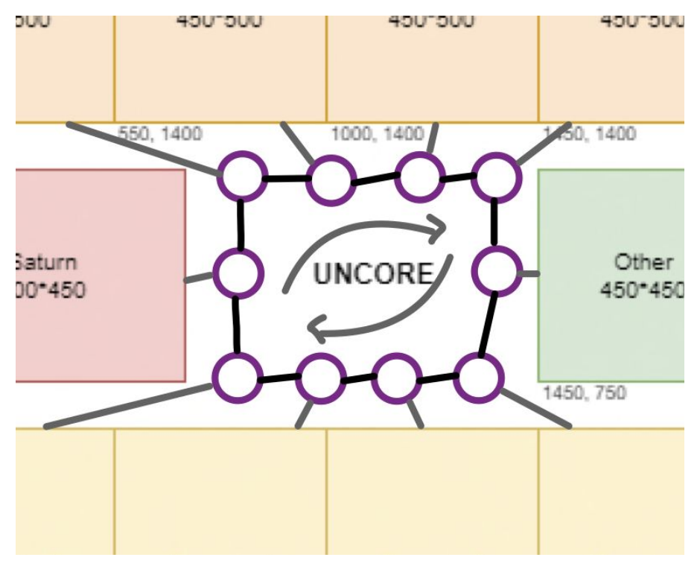
- Rocket Core: 5 stage, in-order
  - configured with 16KB L1D$, 4KB L1I$
- Custom Sparse-Dense Matrix Mult. Accelerator (pictured below)
  - Uses Rocket Custom Coprocessor Interface (RoCC)
- Max operating frequency: **500 MHz**
- Peak Theoretical/Achievable Throughput: **4/2.6 GOPS**





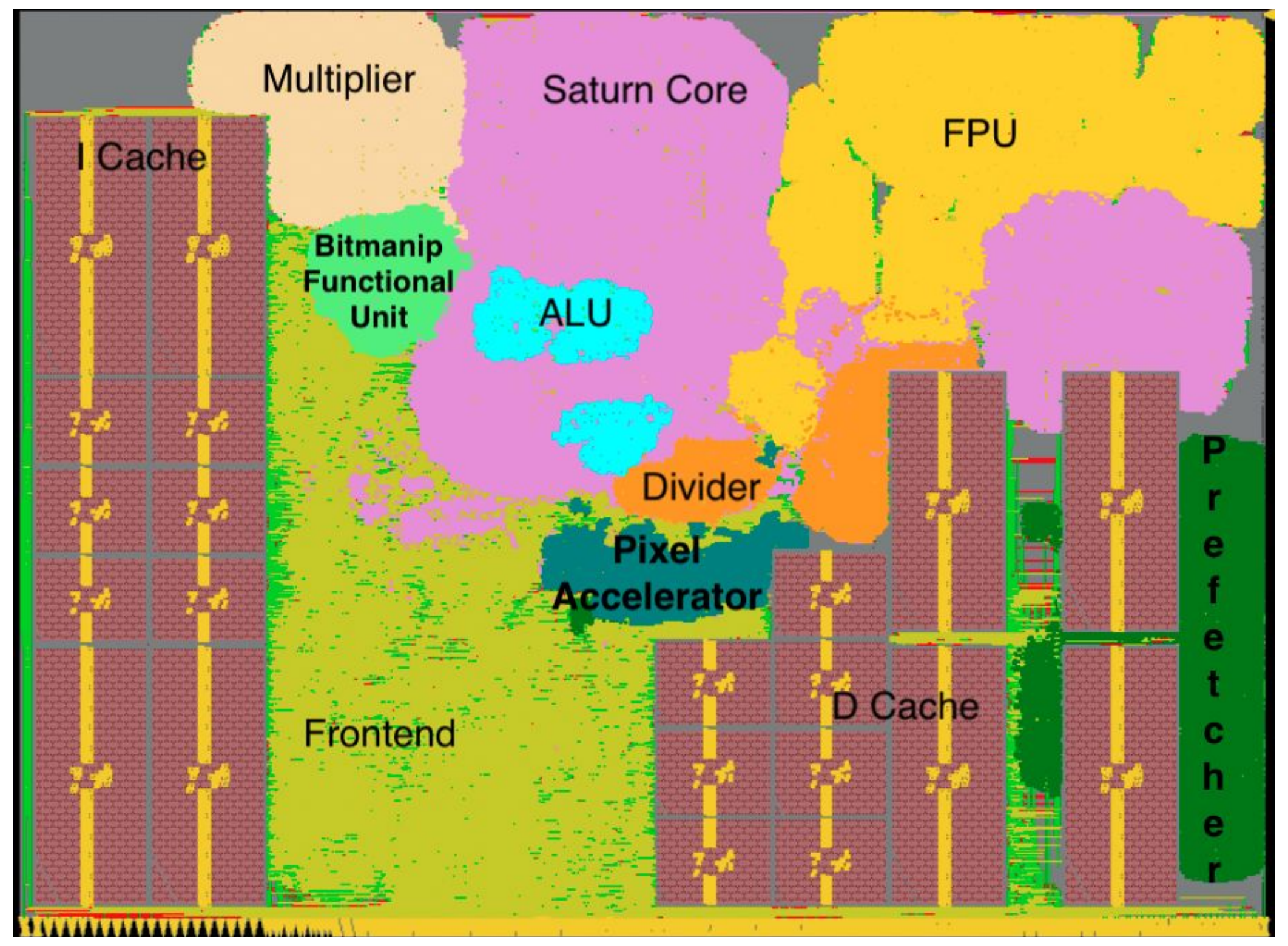
Memory Subsystem

- Network-On-Chip
  - First tapeout of Chipyard's Constellation
  - Rocket cores, Saturn and L2 banks attached to NoC
  - Unidirectional Torus Topology
- Scheduler tiles (L2 $)
  - 8-way set associative
  - 512 KB total L2 capacity divided into 4 scheduler blocks
  - Max operating frequency: **400 MHz**
- Backing scratchpad
  - 16 KB storage space for booting processes
  - Max operating frequency: **400 MHz**



General Purpose Tile

- Saturn Core: 6-stage, dual-issue in-order core
- 16KB I$ & D$; 4 sets & 4 ways, non blocking
- Additional Units
  - Prefetcher
  - Bit-manipulation Functional Unit
  - Pixel Accelerator
- Max operating frequency: **500 MHz**
- Peak throughput: **0.6 GOPS**



Bringup & Next Steps

- PCB Design next semester
- TSI:
  - Configure FPGA and FESVR to work with Bearly ML
  - Involves understanding TSI well to debug
- JTAG:
  - Try connecting with clocks synced
  - Bit bang script that debugs with exposed state machine
  - Involves understanding DTM
- Write software to run neural net with matmul and drivers

Acknowledgements

We would like to thank course instructors Borivoje Nikolic, Ali Niknejad, and Kris Pister, as well as GSIs Jerry Zhao, Bob Zhou, Paul Kwon, Sherwin Afshar, and Zhaokai Liu. Special thanks to Intel for providing this tapeout opportunity and Apple for providing design feedback throughout the semester.