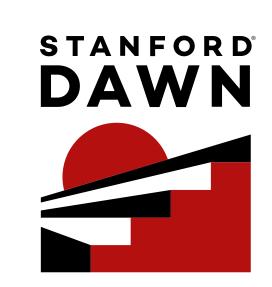


Increasing Dynamism in Plasticine

YaqiZhang yaqiz@stanford.edu

Alexander Rucker acrucker@stanford.edu

Matthew Vilim mvilim@stanford.edu



— Background —

Plasticine is a vector Coarse-Grained Reconfigurable Array:

- 6-stage, 16-lane 32-bit floating point SIMD pipelines
- Distributed 256-kByte memories
- DRAM controllers with tile load and scatter-gather support

Plasticine demonstrated up to 95x speedup vs. an FPGA, and 77 times performance per Watt.

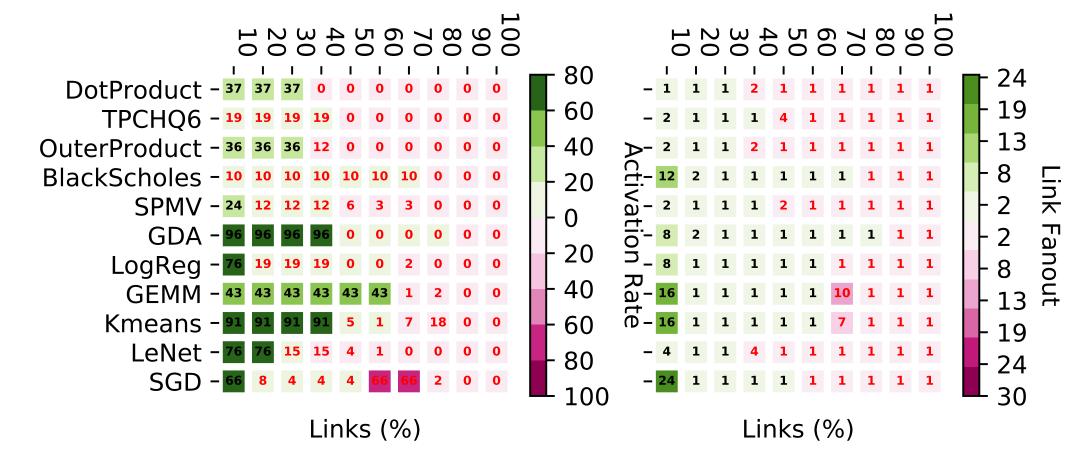
How can we retain Plasticine's performance and efficiency while enabling new applications?

— Compiler & Mapping Flow – Fast Cycle Banking + Placement and Compute Unit Buffering Accurate Control Allocation and Routing Partitioning Simulator VC Allocation Loop Unrolling Integrated with Control Logic **DRAMSim** and Allocation BookSim Host CU \rightarrow CU-B \rightarrow CU-C \rightarrow CU-I \rightarrow Dram Address CU-G ¬ Generator CU DRAM Interface CU CU-A Memory CU Compute CU Scalar Data CU-K - CU-H → CU-D → CU-E → CU-J — Virtual Compute Unit Data-Flow Graph Mapping Compute CU Memory Compute SIMD Pipeline Compute DRAM CU DRAM CU Compute Memory Control Compute DRAM CU Memory

Physical Compute Unit

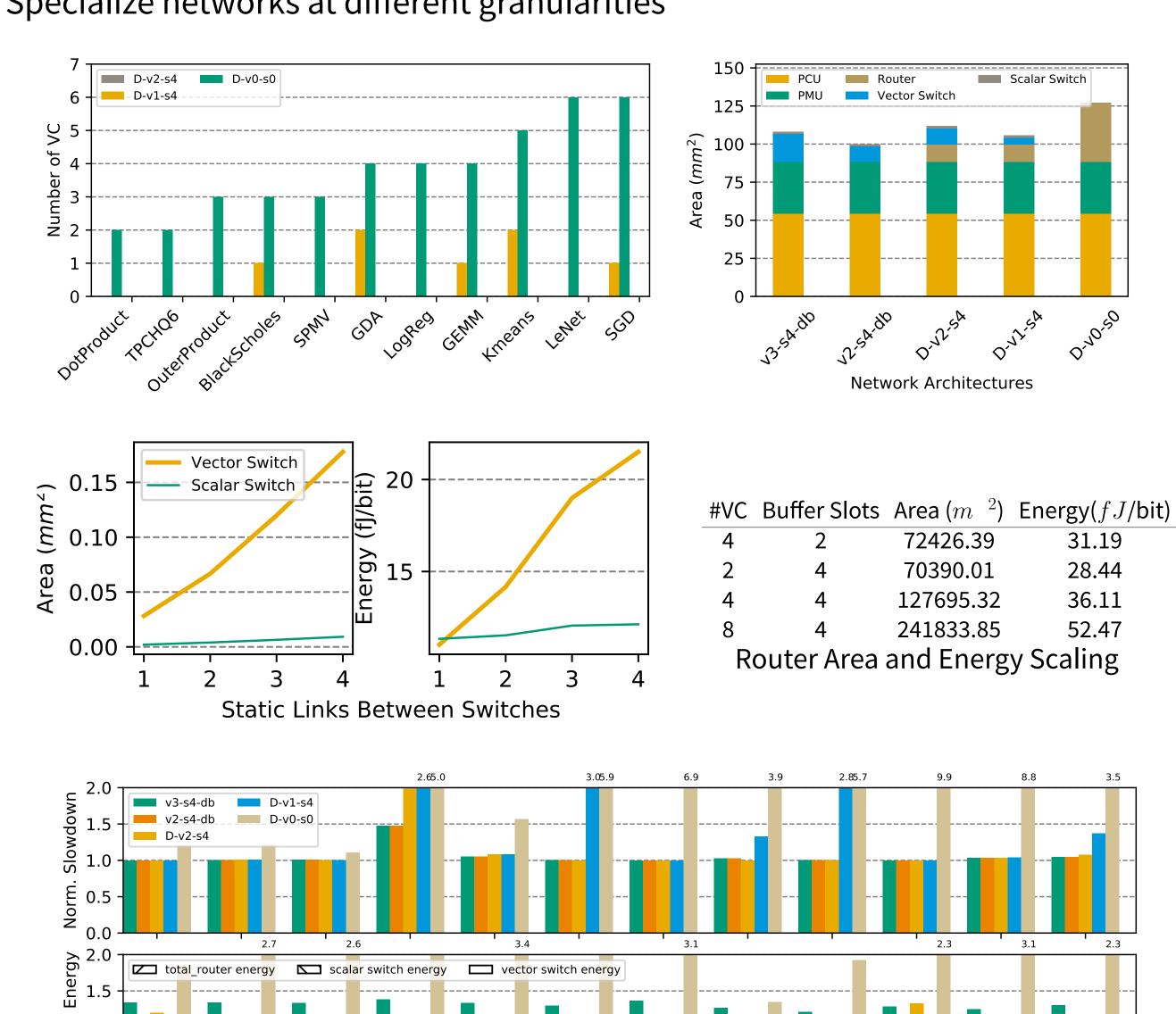
— Hybrid Networks –

Different applications have different link activation rates and fanouts:



How can we improve link utilization?

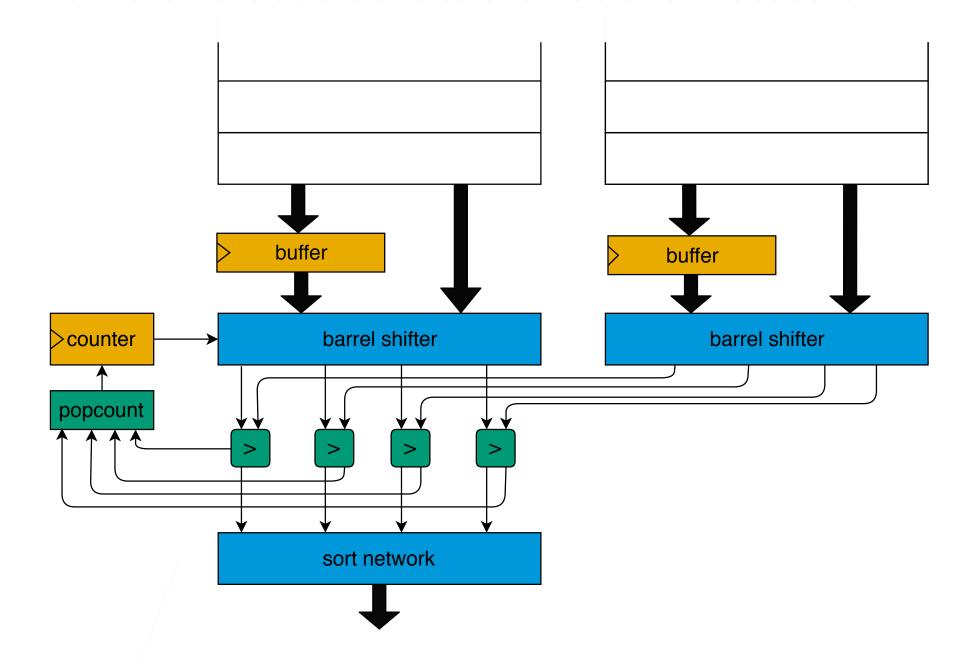
- Use static network for high-bandwidth and broadcast links
- Use dynamic network to encourage link sharing on low-activation links
- Specialize networks at different granularities



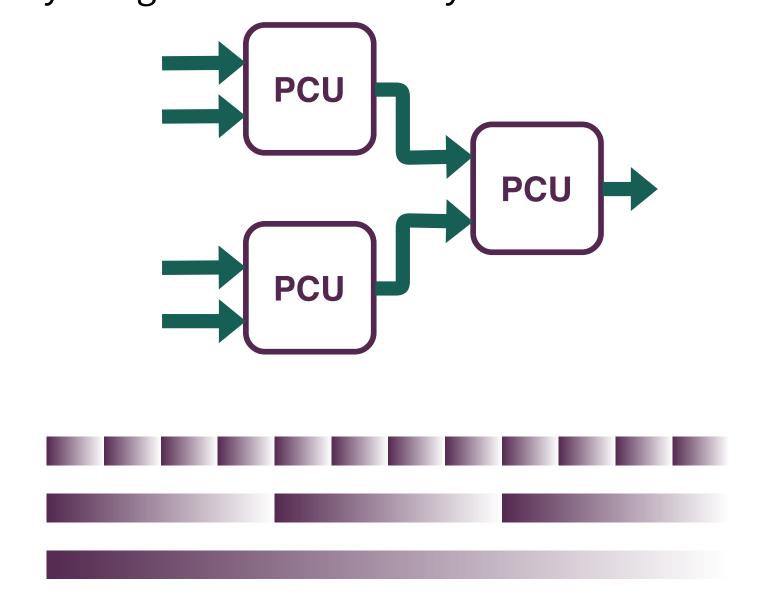
DotProduct TPCHQ6 SuterProduct Scholes SPMV GDA LogReg GEMM Kmeans LeNet SGD Geomean

— Future Work —

Minor hardware additions to the PCU allow fast sort:



Multi-way merge reduces memory traffic:



What's the next class of applications to target?

- Transactional/online applications?
- Streaming data analytics and networking?
- Graph analytics?

What advances will be necessary to target these applications?

- Enhanced support for data-dependent conditionals
- Support for finite state machine-based control
- Tightly integrated parsing support
- Support for more complicated data structures