Deflection Routers

- Pure bufferless deflection routing (CHIPPER) reduces network throughput → reduced application performance
- But, reduced power and area are desirable

Bufferless Routing Reduces Power But Also Degrades Performance

- Bufferless routers eliminate buffers → less static power
- Bufferless routers introduce deflections → higher dynamic power and lower performance at high load

Combine Deflection and Buffering for the Best of Both Worlds

- Key Insight: Starting with pure bufferless deflection routing (CHIPPER), adding a small buffer allows router to buffer some flits and deflect other flits at fine granularity.
  - Deflection rate reduces relative to bufferless routers which deflect all contending flits
  - Buffer is more efficiently used relative to input-buffered routers which buffer all flits

Shortcomings in Prior Bufferless Deflection Routers

1. All contending flits are deflected: high dynamic power and low performance at high load (when many flits contend)
   - Deflection rate in CHIPPER is 28% on average
2. Only one flit can be ejected per cycle: when multiple flits arrive simultaneously, some must be deflected
   - Ejection bottleneck causes deflections in 9% of all cycles in CHIPPER on average (4x4 network)
3. Uncoordinated prioritization unnecessarily deflects: pseudorandom arbitration under Golden Packet leads to priority inversions inside routers

MinBD: Bufferless Deflection Routing

- Side Buffering
  - When flits arrive, perform deflection routing first.
  - Buffer up to one deflected flit in a small “side buffer”.
  - Re-inject side-buffered flits when space is available.

- Dual-width Ejection
  - Replicate ejector module to allow two flits/cycle to eject (captures most demand, eliminates bottleneck)

- Silver-Flit Prioritization
  - Introduce lower Silver Flit priority locally at router
  - Does not interfere with Golden Flit correctness
  - Allows for coordinated deflection arbitration

MinBD: Minimally-Buffered Deflection Routing

- Best energy efficiency of all evaluated designs
- Close to buffered performance for lower cost

Results