Multi-Core Computer Architecture: Storage and Interconnects Week 7

Aronya Baksy

November 2021

1 Introduction to Bufferless Routing (BLESS)

- Buffers increase network bandwidth and throughput
- Router buffers consume lot of power, hence buffered routers are power inefficient

1.1 Bufferless Deflection Routing

- Flits are buffered in pipeline latches and on the network links
- When two packets contend for the same link, one packet is given the desired link and the other one is deflected
- VC allocation logic is now replaced with some flit-ranking and port-prioritization logic
- This logic is applied to each flit in order of ranking
- Each flit is ranked independently. Flits are assigned to productive output port (if possible, or deflected). Flow control is completely local
- Oldest-first ranking: oldest flit has highest priority
- Injection policy: inject flit whenever input port is free
- Absence of deadlocks as each flit always moving. Absence of livelocks with oldest first ranking
- Advantages of BLESS:
  - No buffers, no local flow control
  - No deadlocks and livelocks
  - Adaptable: deflect packets around congested links
  - Less area needed for router
- Disadvantages of BLESS;
  - Increased latency, increased buffering at receiver for reassembly of out-of-order flits
  - Reduced bandwidth
  - Header information has to be included in each flit (not only head flit)
  - Oldest-first ranking is complex

1.2 Livelock Freedom in BLESS

- All flits are timestamped. Oldest flits get productive ports.
- Router needs to sort flits by age. Sorting is complex and long latency (3 comparator stages for 4 flit)
- The allocation of flits on a priority basis is a sequential procedure. First high ranking flit is allocated, then one by one lower ranked flits are allocated
- Overall, this causes BLESS to have longer critical path than buffered router

1.2.1 Golden Flit Concept

- Key idea: no need to maintain total ordering between all flits. Just pick best flit vs rest flits at each time and allocate it to the profitable output
- 2 input: Pick a winning flit (if both input flits are not golden then pick at random). Steer the winning flit to the appropriate direction
- In this case, a parallel flit allocation mechanism can be used.
- There are 2 stages. Each stage is a 2 input router as shown. The packets are swapped if the higher priority one needs to move to another port, otherwise no swap.

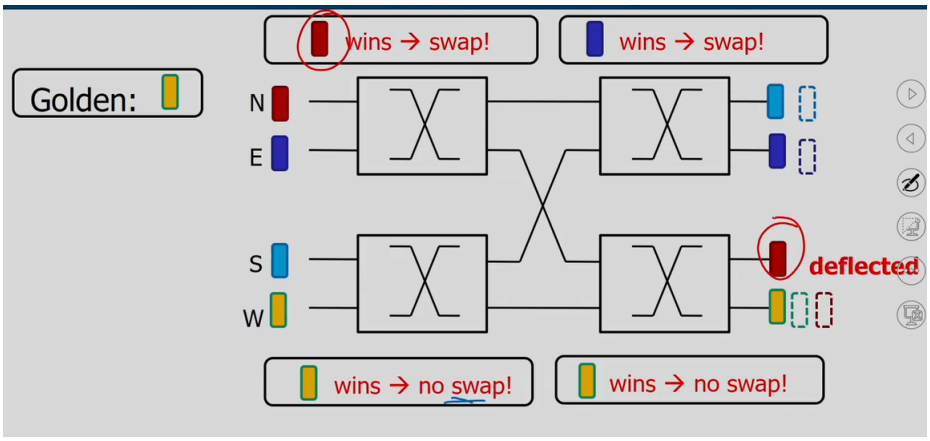


Figure 1: Permutation Deflection Unit

1.3 CHIPPER Router Architecture

- Published by Chris Fallin et. al. in HPCA 2011
- Early injection and ejection stage before port allocation
- Idea: If all inputs are busy then don't inject new flits.
- Drawbacks of CHIPPER: high deflection rate when traffic is
- Performance issues of CHIPPER:
  - Due to lack of buffers, any link contentions causes deflection
  - Ejection bottleneck: Only one flit can be ejected per router per cycle (hence simultaneous arrival causes deflection)
  - Practical deflection arbiters (golden flit) cause unnecessary deflections
  - These are solved in MinBD
- Performance results: 36.2% less area than buffered, 29.1% smaller critical path than BLESS

2 Minimally Buffered Deflection Router (MinBD)

2.1 Side-Buffering

- Step 1: remove up to 1 deflected flit per cycle from the output of the parallel port allocation unit
- Step 2: Buffer this removed flit in a small FIFO side buffer
- Step 3: Re-inject this flit from side buffer into the pipeline when a slot is available
- **Dual-Width Ejection**: Allow 2 flits to be ejected in the same cycle, using 2 ejection units in the pipeline (instead of 1 before). This solves the ejection bottleneck.

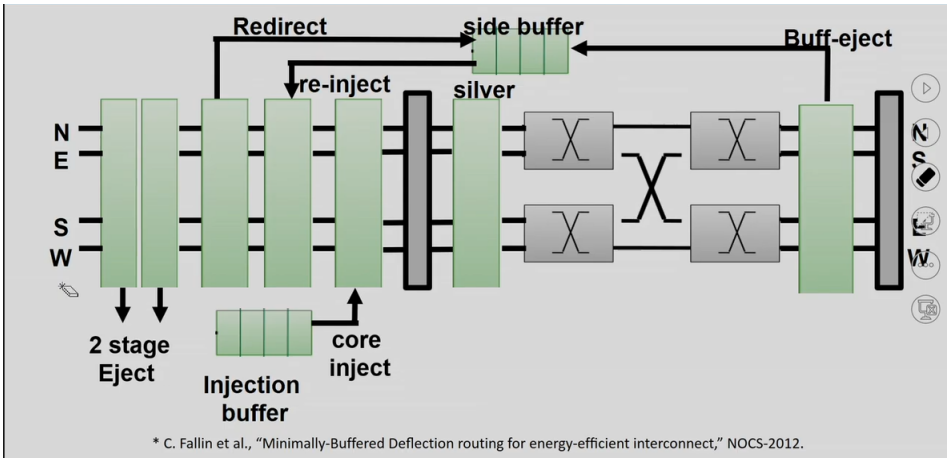


Figure 2: Minimally Buffered Deflection Router (MinBD)

2.2 Improving Deflection Arbitration

- Unncessary deflections are caused by lack of coordination (random choice when all flits have same priority)
- Add one priority level to ensure that atleast 1 flit is not deflected in each cycle
- Highest priority is golden flit. Selected by static round-robin schedule, and full network contains only one golden flit. (correctness)
- Next highest priority is silver flit. One silver flit per router, chosen in a pseudo-random manner. (performance)

2.3 Limitations of MinBD

- 2 separate ejection stages and dual ejection port (more power, more area)
- Pre-emption takes care of stravation in side buffer only. Core buffer starvation is not addressed
- Random selection of silver flit does not ensure progress (a silver flit in one router, may not be silver in next)

3 Deflection-Based Adaptive Router (DeBAR)

- John Jose et. al, DATE 2013
- Stage 1: Hybrid Ejection Unit, Flit Pre-Emption Unit, Dual Injection Unit
- Stage 2: Priority Fixer Unit, Quadrant Routing Unit, permuation Deflection Network, Buffer Ejection Unit

3.1 Hybrid Ejection Unit

- Analyze whether a flit is to be ejected or not using a single ejection port
- If 2 flits are to be ejected, then one is ejected to the output, the other one is buffered in the ejection bank of side buffer.
- If a flit is already waiting in the ejection bank of side buffer, it is replaced whenever a new flit is to be ejected. The old flit already in the EB is moved out and the new ejected flit takes its place
- If 2 flits are to be ejected but EB is already full, then empty the EB by ejecting its content, then move one of the new flits to the EB and allow the other one to pass through (it will get deflected later, but this happens only 0.025% of the time)

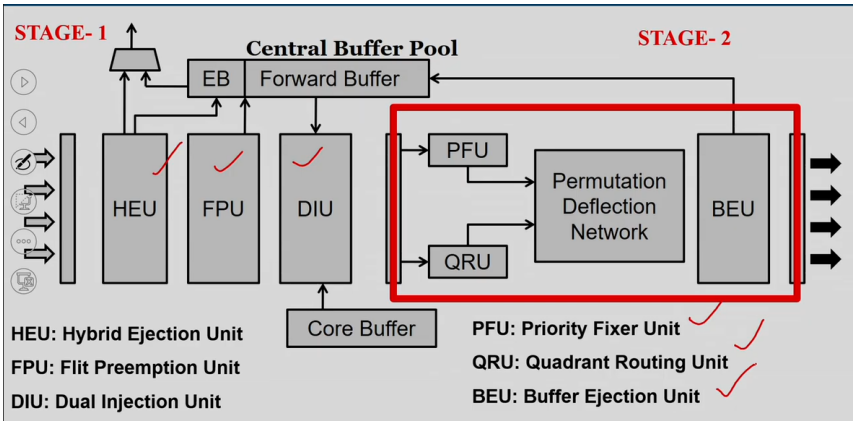


Figure 3: Deflection-based Adaptive Router (DeBAR)

3.2 Dual Injection Port

- Solves the problem of starvation in core buffer
- Uses odd-even based priority.
- **One free channel**: In odd clock cycles, flits from side buffer are injected into free channel. In even cycles, flits from core buffer are injected
- **2 free channel**: one flit from core and one from side buffer are injected

3.3 Flit Pre-Emption Unit

- If the input channels are all occupied then both core and side buffers starve. In this case, the Flit Pre-Emption UNit helps
- In this case, define a reinjection interval for side buffer and a core-injection interval for core buffer (RII and CII).
- RII and CII are respectively, the interval for which a flit in the side/core buffer can wait before being injected
- The FPU pre-empts one of the channels, buffers its flit in the side buffer and pushes the head of the side buffer to the empty channel

3.4 Priority Fixer Unit

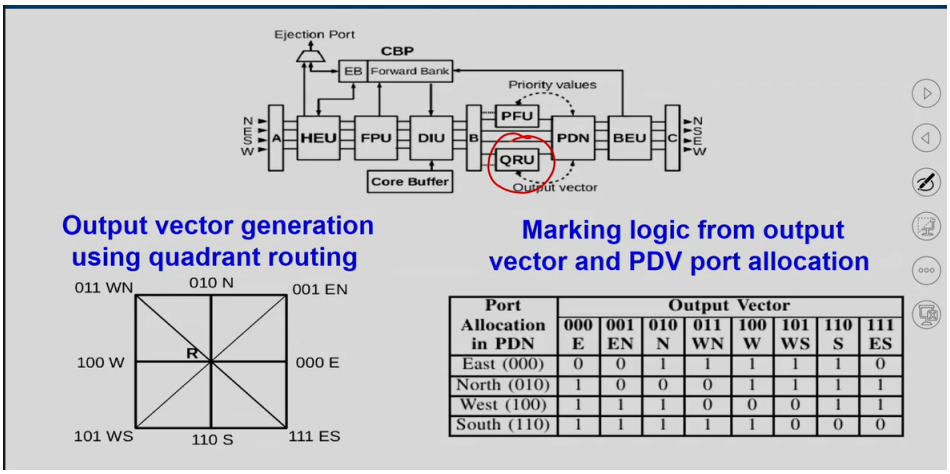
- Compute hops to destination for each flit and assign priority value to each flit based on that
- Hops to destination is computed using Manhattan distance from current posn to dest

3.5 Quadrant Routing and Marking

- The 8 possible directions of the mesh are given a 3 bit output vector
- This vector represents the relative position of the current router to the destination fo that flit
- If port allocated is not productive for current flit, then flit is **marked** by the PDN. Marking allows next port to take an informed decision

3.6 Buffer Ejection Unit

- If mark bit is 0, then allow that flit to pass through
- If mark is 1, then move flit to side buffer



Port Allocation in PDN	Output Vector 000 E	001 EN	010 N	011 WN	100 W	101 WS	110 S	111 ES
East (000)	0	0	1	1	1	1	1	0
North (010)	1	0	0	0	1	1	1	1
West (100)	1	1	1	0	0	0	1	1
South (110)	1	1	1	1	1	0	0	0

Figure 4: Quadrant Routing and Marking

3.7 Summary of DeBAR

- Dual ejection support without 2 ejection port and 2-stage ejection unit
- Ensure fairness in progress of flits between core and side buffers
- Priority computation and output vector generation in parallel
- PDN does port allocation and marking
- Non-uniform side buffer size. Central routers have 4 flit space, edge routers have 2

3.8 Results

- DeBAR accomodates more packets in network before showing saturation for many different network traffic patterns (avg latency benchmark)
- DeBAR has less deflection rate for different network traffic patterns than MinBD
- DeBAR has 17% less latency in stage 1 than MinBD. Stage 2 is approx same for both.
- 18% reduction in channel wiring due to 1 ejection port in DeBAR vs MinBD

3.9 Drawbacks

- Channel wastage when Buffer Ejection Unit removes the non-productive flits when there may be flits looking for that channel (18%)
- Internal movement of flits from core to side buffer (11%) or side buffer to side buffer (22%)
- Older flit penalization

4 Smart Late Injection Deflection Router (SLIDER)

- Bhawna Naik et. al. ICCD 2013
- Key ideas:
  - Parallelizing independent operations
  - Selective pre-emption
  - Smart late injection

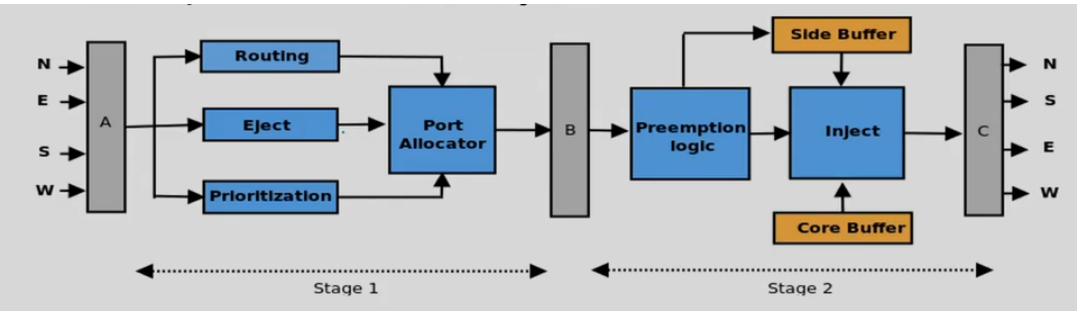


Figure 5: Smart Late Injection Deflection Router (SLIDER)

4.1 Smart Late Injection

- Flit is injected only if its productive port is available. **Late** injection allows this to happen. This is called restricted injection
- because of restricted injection, the buffers will fill up. When core buffer is occupied more than half, then move to non-restricted injection
- Non-restricted injection, move the head of core buffer to the empty port even though it is not productive
- This is shown to reduce deflection rate and packet latency over DeBAR
- 25% piepline latency reduction, 19% power reduction, 3.5% area reduction over DeBAR