Lecture 9-10
Instruction Supply: Branch Prediction

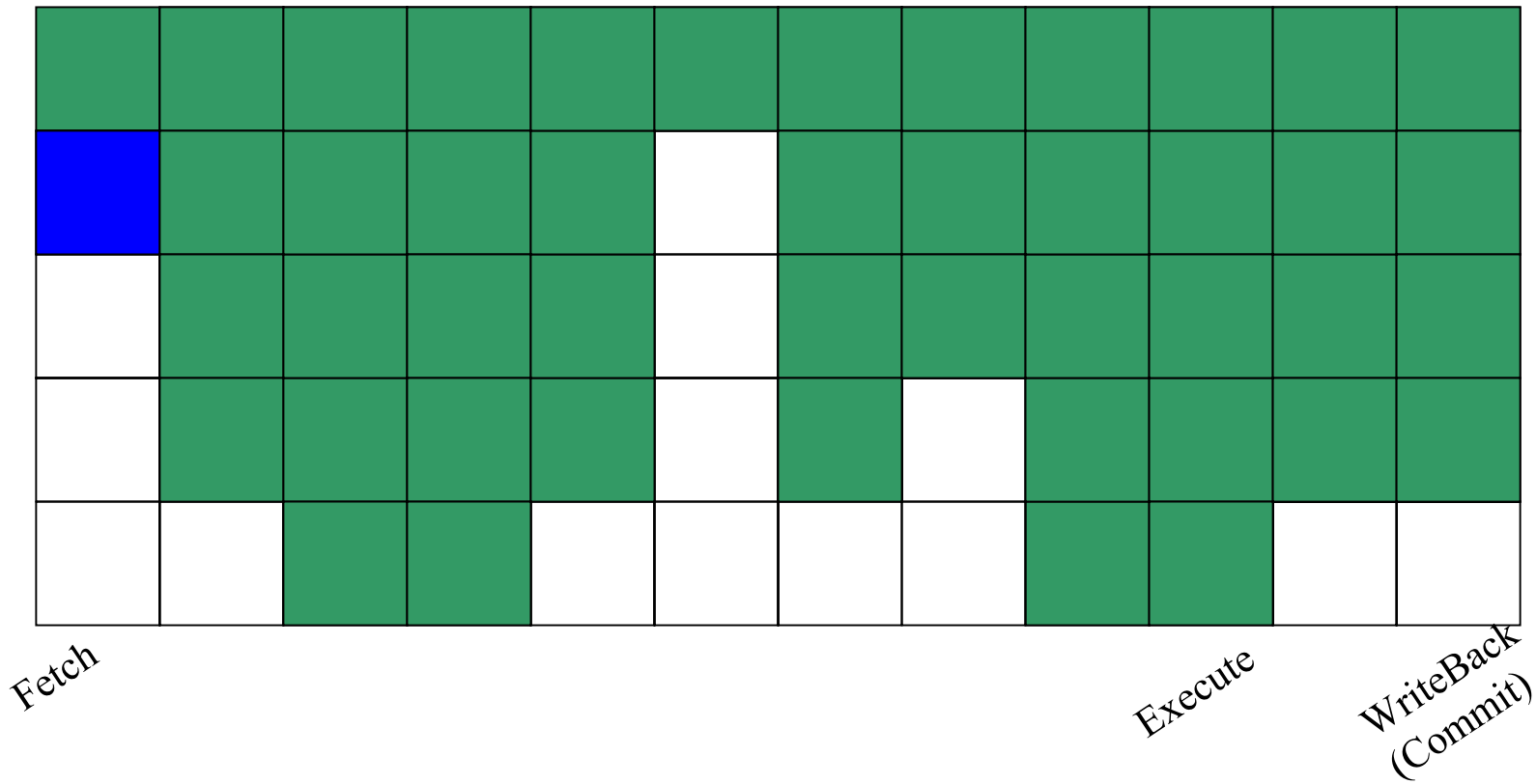
Outline

- The effect of branches on performance
- Front End Design
- Branch Predictors
  - Static
  - Simple dynamic
  - Correlation based

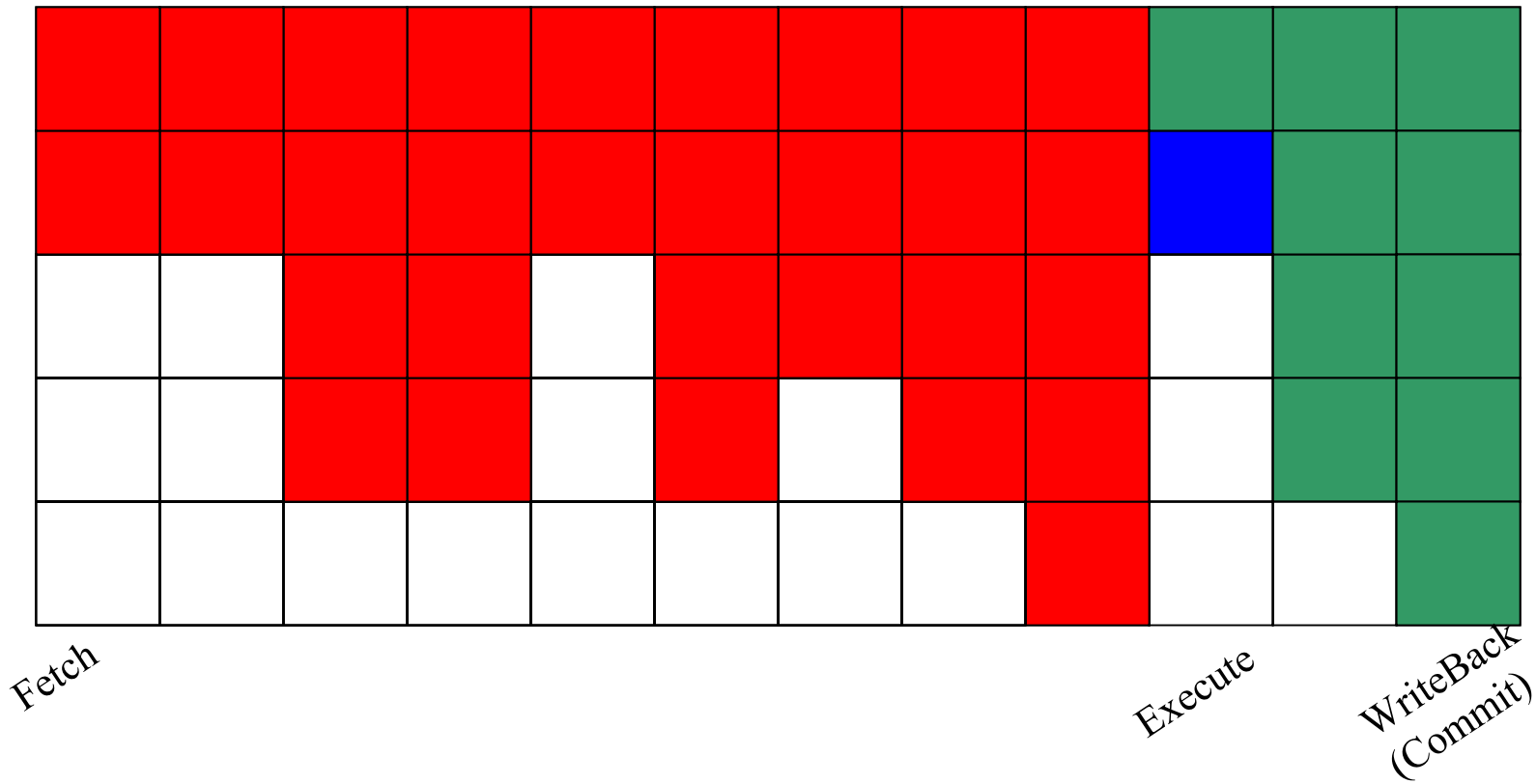
Maximizing Instruction Supply

- Reducing branch mispredictions
- Reducing fetch misses, e.g., icache misses
- Reducing partial fetches

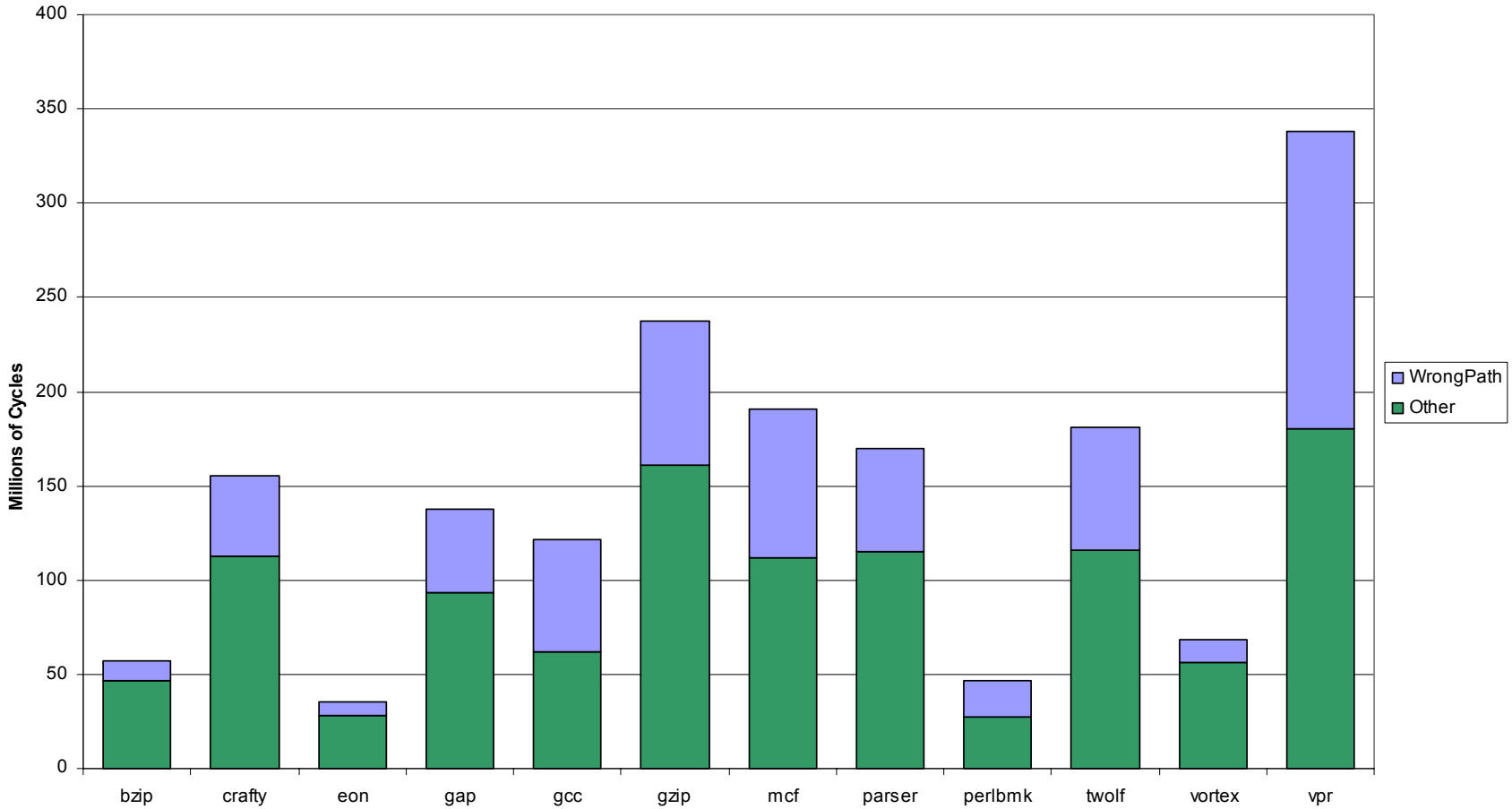
The effect of branches



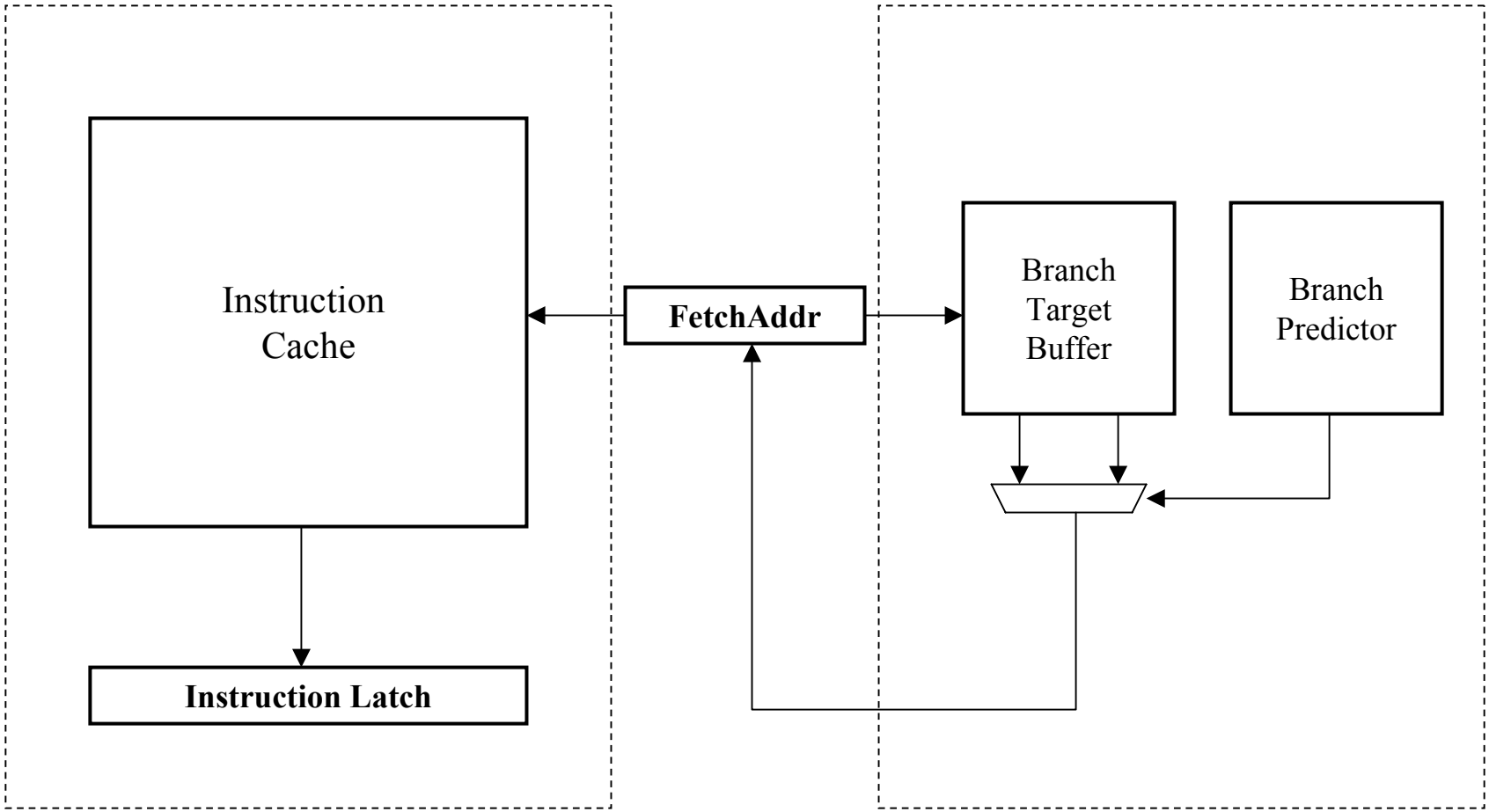
The effect of branches



Impact of Branch Mispredictions
[B. Fahs et al, 2001]

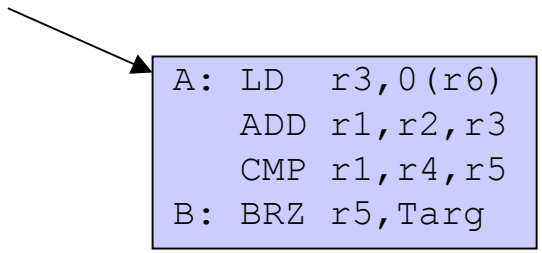


Front End Design



Branch Target Buffer

- An incorrect target is called a misfetch
- BTB associates FetchAddr with NextFetchAddrs



BTB associates A with B+4 and Targ

Typical BTB Designs

- Must be single cycle (why?)
- Set-associativity is desirable (why?)
- Alpha 21264 : 1K Entries??
- P6 Family: 512 Entries
- PPC 620: 256 Entries

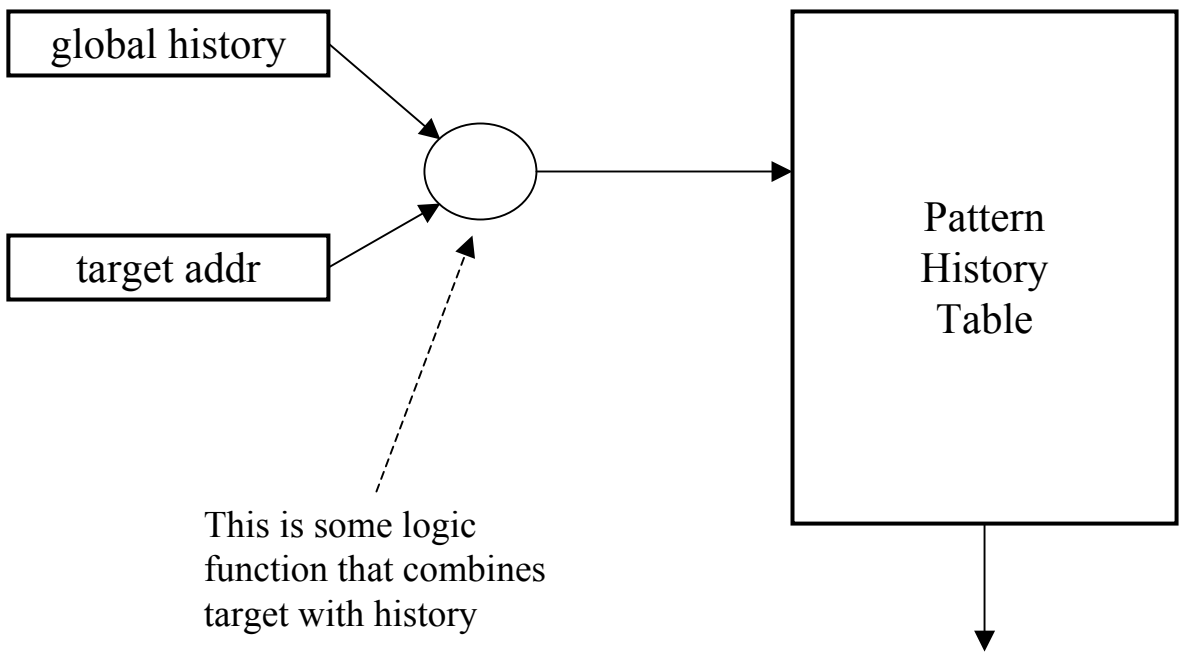
Static Branch Prediction

- Can be based on branch opcode/offset
  - Backwards Taken/Forward Not Taken
- Can be based on compiler hints
  - Often based on *profiling*
- Where are the static predictions stored??

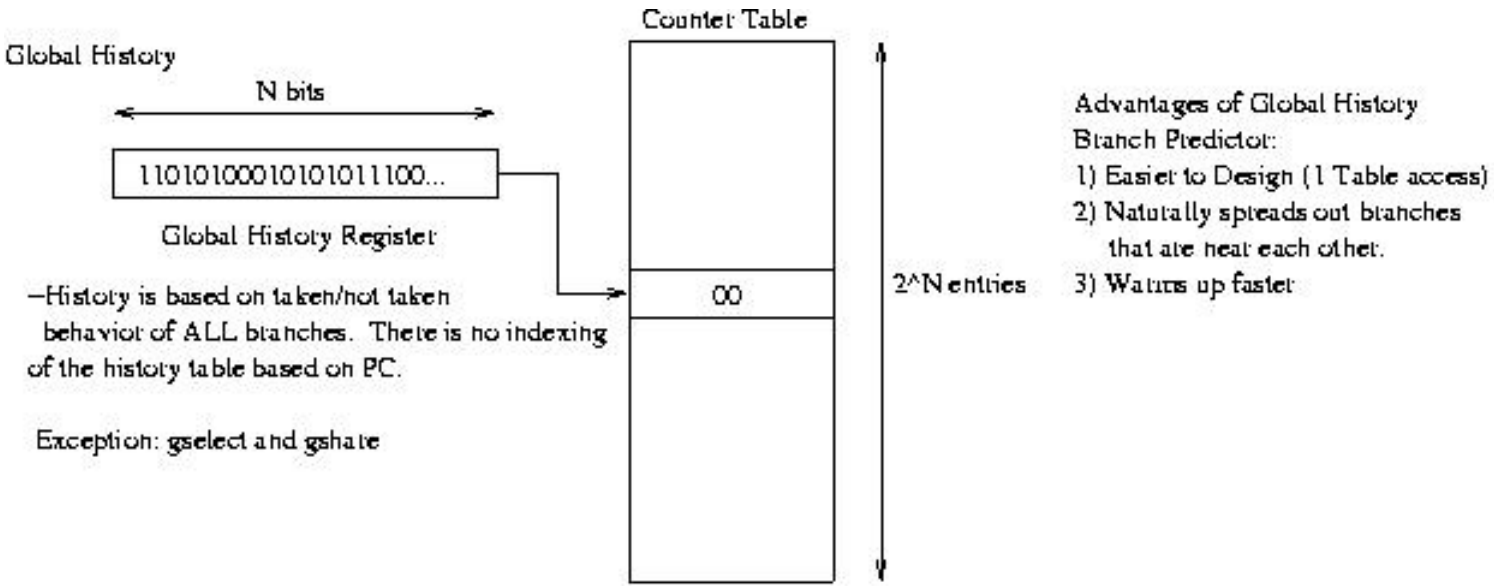
Dynamic Branch Prediction

- Single level (simple counters)
- Global Correlation
  - GAs
  - gshare
- Local Correlation
  - PAs
- Hybrid Predictors

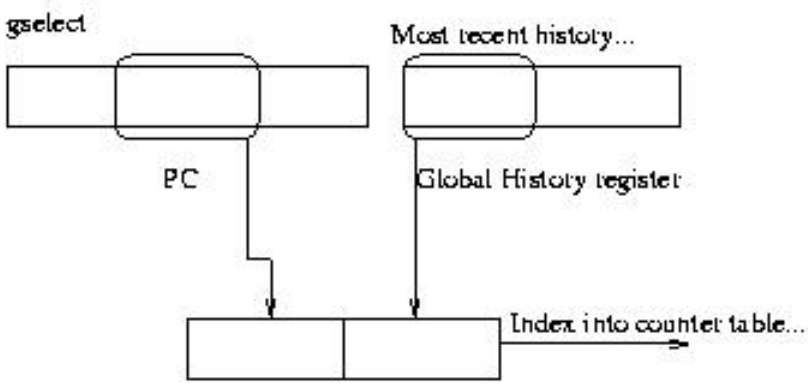
Global History

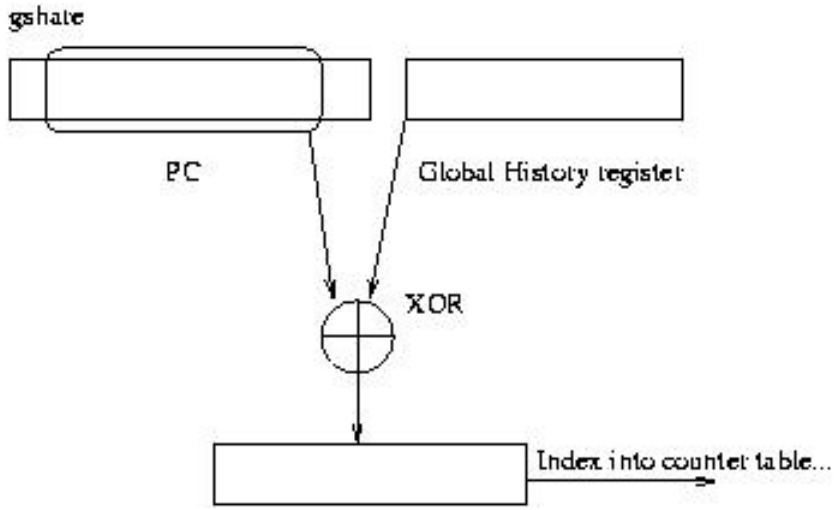


Global History

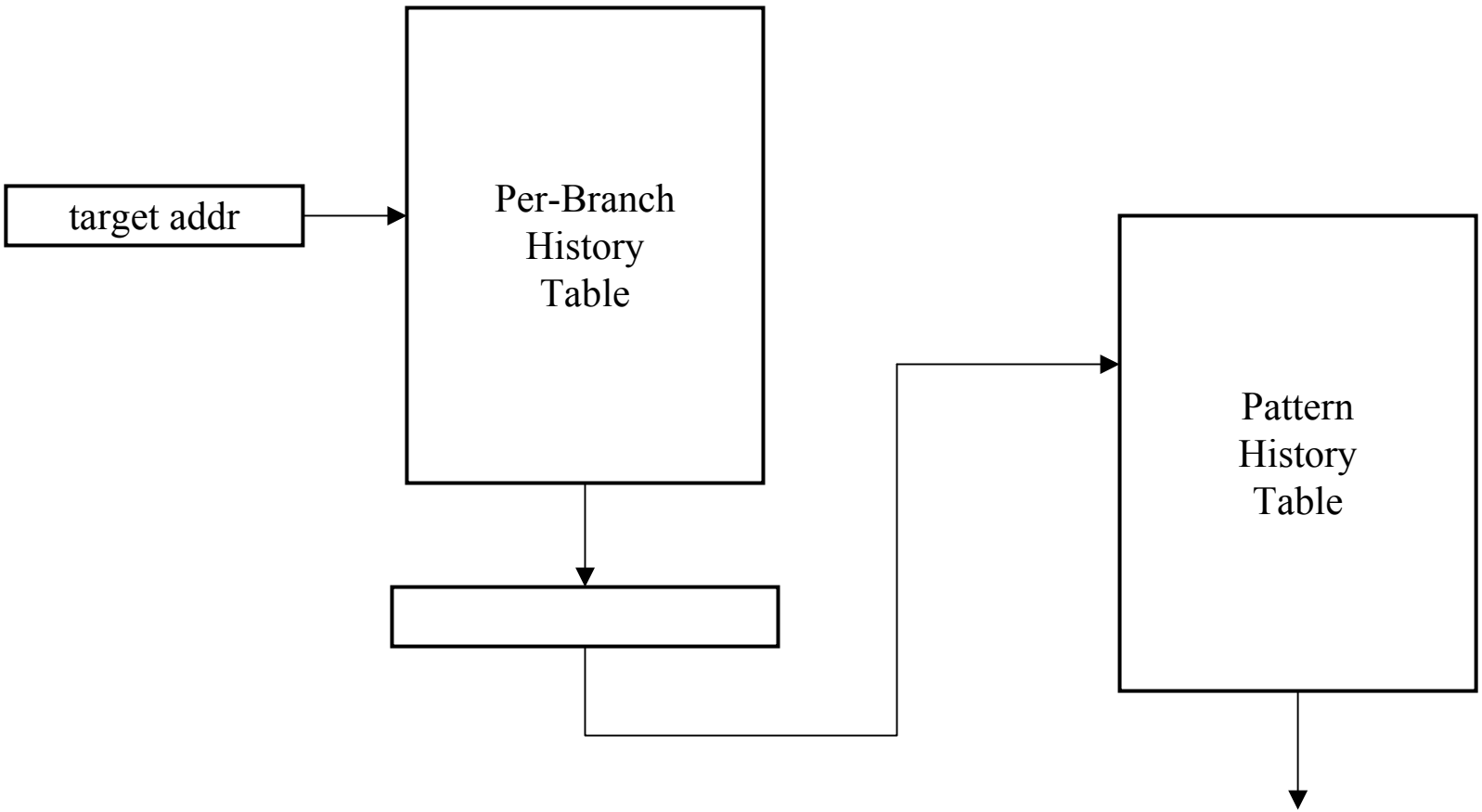


gshare and gselect

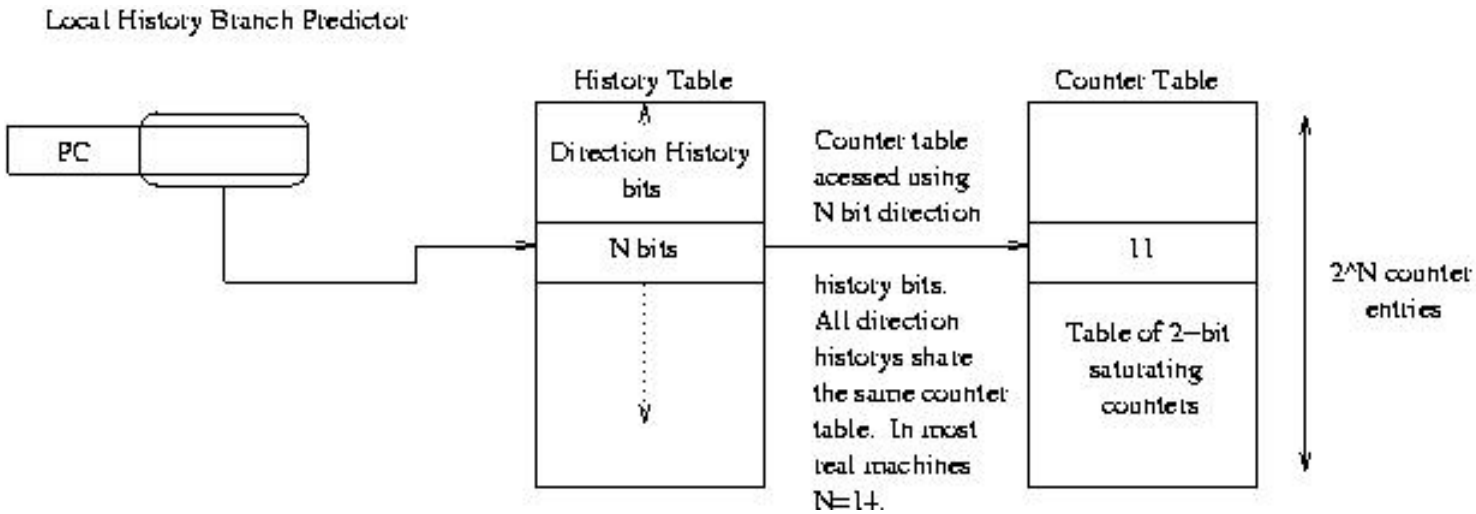




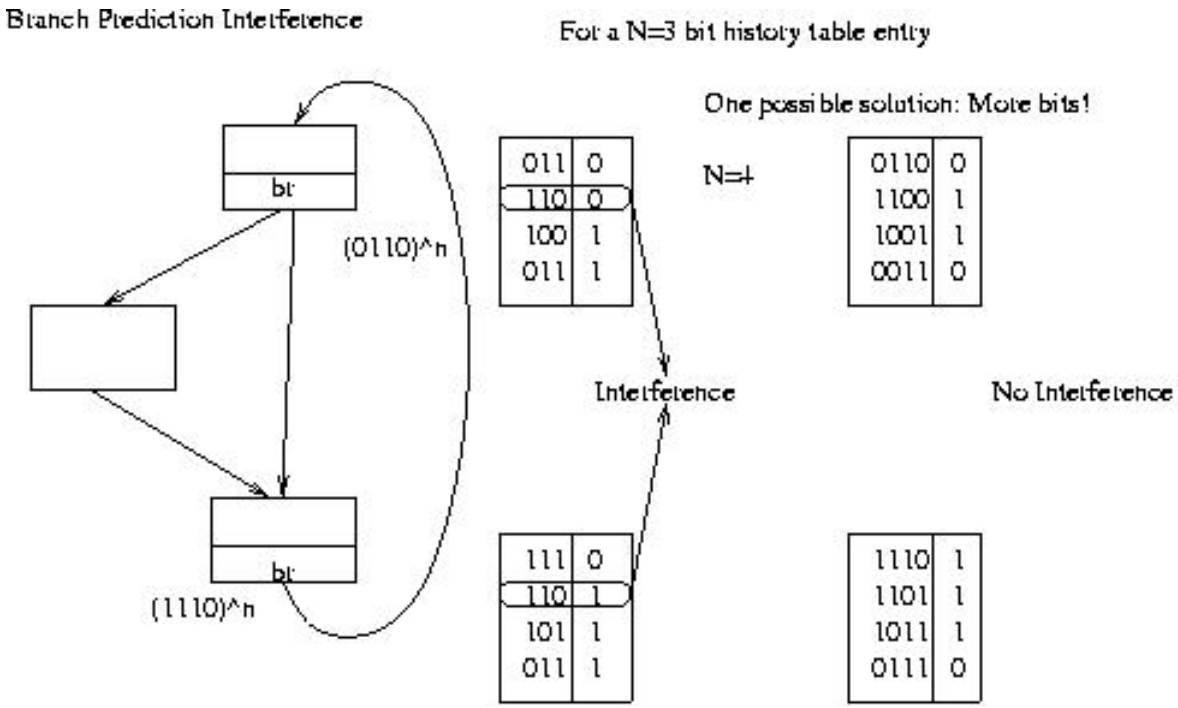
Local History



Local History



Local History Branch Prediction Interference



Questions:

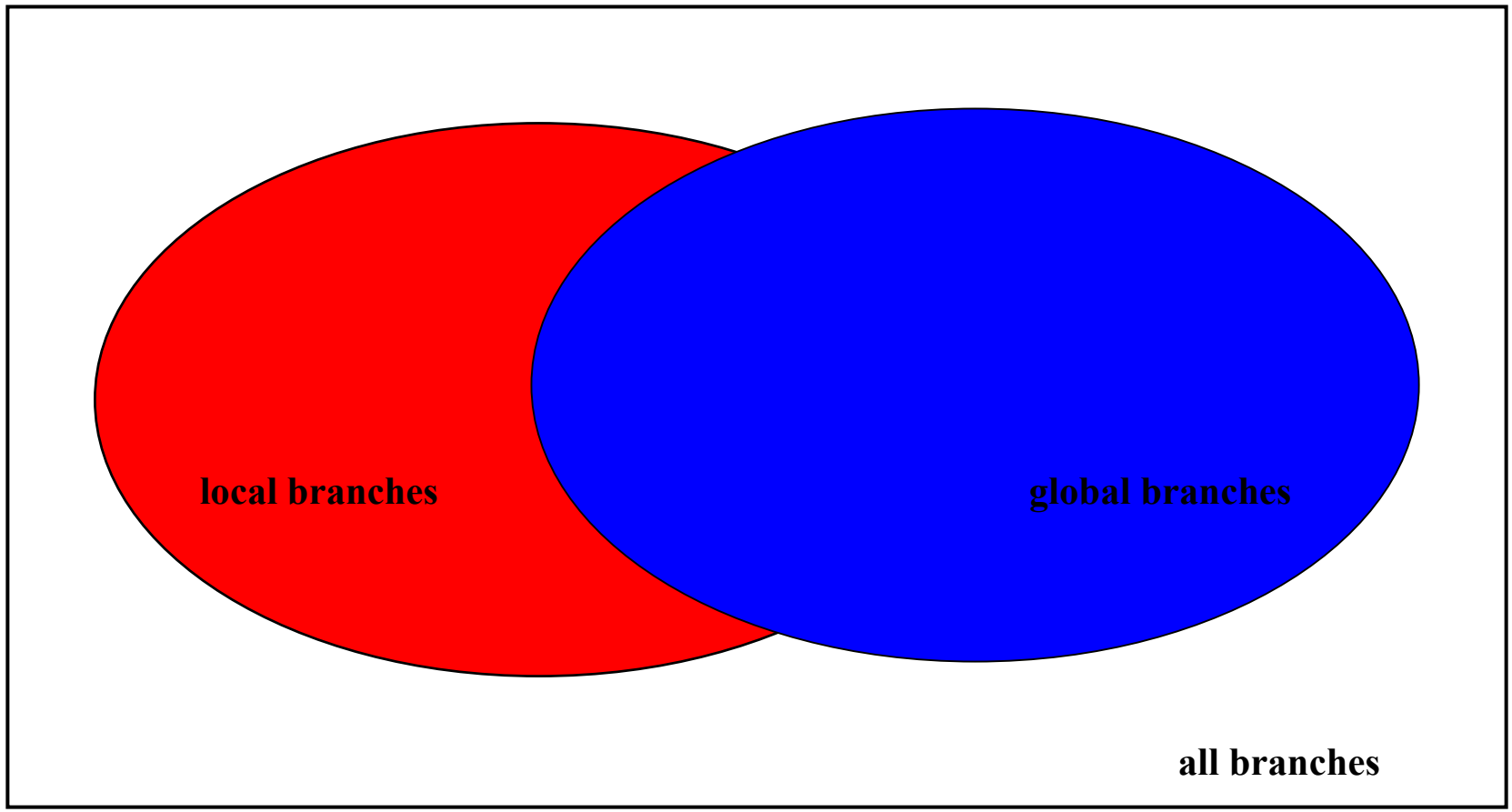
- What happens to the branch predictors on a misprediction? In other word, what has to be repaired because of the mispredict?

- The local predictor requires two table accesses – how can we build it such that given an address, a prediction can be made with one table access?

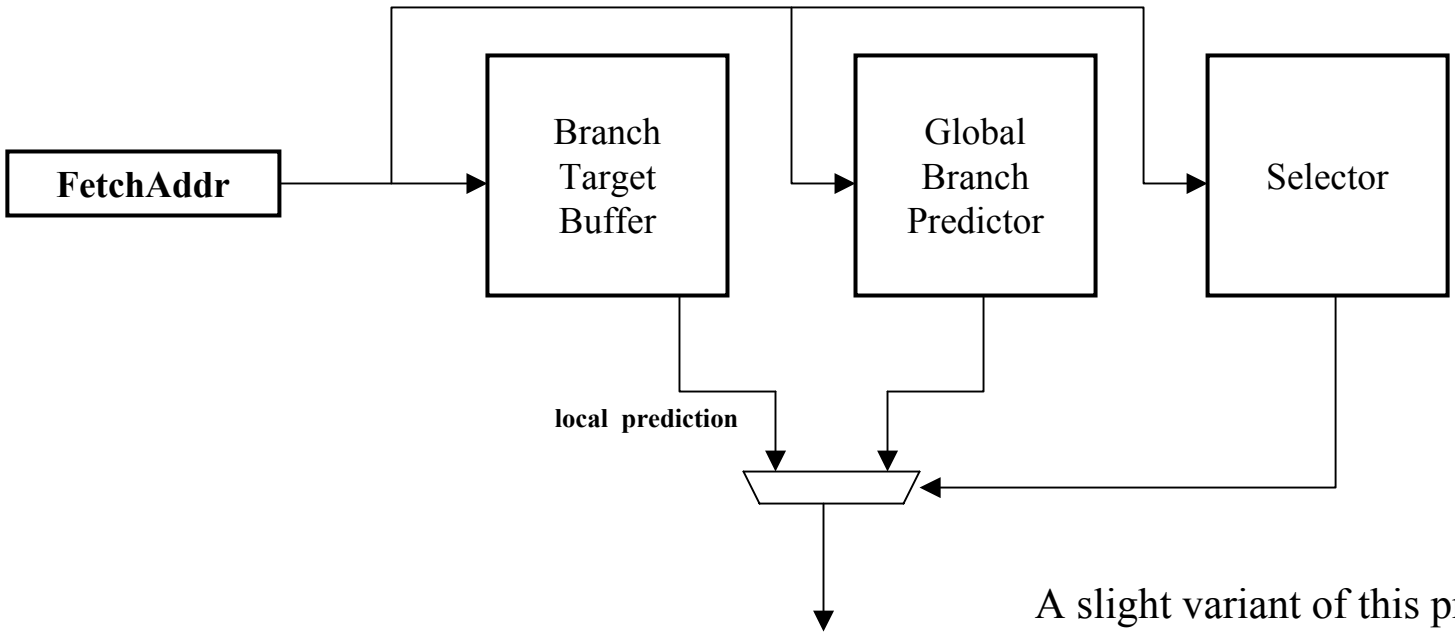
Outline

- Hybrid Predictors
- Repair of predictors
- Interference
- Other types of branches
- Beyond the current state of the art

Local vs Global Correlation



Hybrid Predictor Design



A slight variant of this predictor appeared in the Alpha 21264

Update and Repair of predictors

- Update of branch predictor state
  - when do we update history?
  - when do we update counter?
- Repair: what do on a misprediction
  - how do we repair history?
  - how do we repair counter?

Interference

- Two or more branches that are not correlated share the same PHT entry.
- Two types of interference
  - negative
  - positive
- What to about it:
  - filtering
  - agree predictor [in the reading list]
  - and half a dozen other ideas…

Other types of branches

- Return from Subroutine (RET)
  - Account for about 10% of all branches
  - Mechanism: Return Address Stack
- Indirect Jumps (JMP Rx)
  - Account for about 3% of all branches
  - Mechanism: adding history to BTB index

Future outlook

- Correlation to data values
- Correlation to specific branches
- Control-independence
- Predicting predictability