Material from Section 4.3

This set under construction.

Outline

- Branch Prediction Overview
- One-Level Predictor
- Two-Level Correlating Predictor
- Other topics to be added.
- Sample Problems
Branch and Target Prediction

Motivation

Branches occur frequently in code.

At best, one cycle of branch delay; more with dependencies.

Therefore, impact on CPI is large.

Techniques

*Branch Prediction:*
Predict outcome of branch. (Taken or not taken.)

*Branch Target Prediction:*
Predict branch or other CTI’s target address.

*Branch Folding:*
Replace branch or other CTI with target instruction.
Branch Prediction

Methods Covered

One-level predictor, a.k.a. bimodal.

Two-Level Predictors

GAg, a.k.a. Global History.

gshare.

Local History, a.k.a. PAg.
Idea: Predict using past behavior.

Example:

```assembly
LOOP:
lw r1, 0(r2) ! Read random number, either 0 or 1.
addi r2, r2, #4
slt r6, r2, r7
beqz r1, SKIP
addi r3, r3, #1
SKIP:
bneq r6, LOOP ! Loop executes 100 iterations.
nop
```

Second branch, `bneq`, taken 99 out of 100 executions.

Pattern for `bneq`: T T T ... NT T T T

First branch shows no pattern.
SPEC89 benchmarks on IBM POWER (predecessor to PowerPC).
FIGURE 4.15 Prediction accuracy of a 4096-entry two-bit prediction buffer versus an infinite buffer for the SPEC89 benchmarks.
Branch Prediction Terminology

*Outcome:* [of a branch instruction execution].
The outcome of the execution of a branch instruction.

*T:*
A taken branch.

*NT: or N*
A branch that is not taken.

*Prediction:* [made by branch prediction hardware].
The predicted outcome of a branch.

*Misprediction:*
An incorrectly predicted outcome.

*Prediction Accuracy:* [of a branch prediction scheme].
The number of correct predictions divided by the number of predictions.
Branch Prediction Terminology (Continued)

*Speculative Execution:*
The execution of instructions following a predicted branch.

*Misprediction Recovery:*
Undoing the effect of speculatively executed instructions ... ... and re-starting instruction fetch at the correct address.
One-Level Branch Predictor

Idea: maintain a *branch history* for each branch instruction.

*Branch History:*
Information about past behavior of the branch.

Branch histories stored in a *branch history table.*

Often, branch history is sort of number of times branch taken... ... minus number of times not taken.

Other types of history possible.

Branch history read to make a prediction.

Branch history updated when branch outcome known.
Branch History Counter

If a counter used, branch history incremented when branch taken . . .
. . . and decremented when branch not taken.

Symbol $n$ denotes number of bits for branch history.

To save space and for performance reasons . . .
. . . branch history limited to a few bits, usually $n = 2$.

Branch history updated using a saturating counter.

A saturating counter is an arithmetic unit that can add or subtract one . . .
. . . in which $x + 1 \rightarrow x + 1$ for $x \in [0, 2^n - 2]$ . . .
. . . $x - 1 \rightarrow x - 1$ for $x \in [1, 2^n - 1]$ . . .
. . . $(2^n - 1) + 1 \rightarrow 2^n - 1$ . . .
. . . and $0 - 1 \rightarrow 0$.

For an $n$-bit counter, predict taken if counter $> 2^{n-1}$.
One-Level Branch Predictor Hardware

Illustrated for Chapter-3 DLX implementation . . .
. . . even though prediction not very useful.

Branch Prediction Steps

1: Predict.

Read branch history, available in ID.

2: Determine Branch Outcome

Execute predicted branch in usual way.

3: Recover (If necessary.)

Undo effect of speculatively executing instructions, start fetching from correct path.

4: Update Branch History
Branch History Table

Stores branch histories,

Implemented using a memory device.

Address (called index) is *hash* of branch address (PC).

For $2^m$-entry BHT, hash is $m$ lowest bits of branch PC *skipping alignment*.

<table>
<thead>
<tr>
<th>BHT Addr</th>
<th>Align.</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>m+2</td>
</tr>
<tr>
<td>m+1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Data input and output of BHT is branch history.
Sample Local History

Outcomes for individual branches, categorized by pattern, sorted by frequency.

Branches running TeX text formatter compiled for SPARC (Solaris).

<table>
<thead>
<tr>
<th>% Patterns</th>
<th># Branches</th>
<th>gshre local corr</th>
<th>Local History</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>fe7f</td>
<td>0.0004</td>
<td>1397 0.912 0.916 0.896 TTTTTTTNNTTTTTTTT 0</td>
</tr>
<tr>
<td>1</td>
<td>ff3f</td>
<td>0.0004</td>
<td>1323 0.924 0.909 0.900 TTTTTTNNTTTTTTTTT 0</td>
</tr>
<tr>
<td>2</td>
<td>fcff</td>
<td>0.0004</td>
<td>1317 0.949 0.939 0.948 TTTTTTTTTNNTTTTTTTT 0</td>
</tr>
<tr>
<td>3</td>
<td>ff9f</td>
<td>0.0003</td>
<td>1245 0.910 0.905 0.898 TTTTTNNNTTTTTTTTTTTT 0</td>
</tr>
<tr>
<td>4</td>
<td>f9ff</td>
<td>0.0003</td>
<td>1235 0.955 0.950 0.955 TTTTTTTTTTNTTTTTTTTTT 0</td>
</tr>
<tr>
<td>5</td>
<td>ffcf</td>
<td>0.0003</td>
<td>1188 0.926 0.921 0.923 TTTTNNNTTTTTTTTTTTTTT 0</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>0.0003</td>
<td>1163 0.873 0.829 0.854 NNNNNNTNNNNNNNNNNNNN 0</td>
</tr>
<tr>
<td>7</td>
<td>180</td>
<td>0.0003</td>
<td>1159 0.955 0.914 0.926 NNNNNNNNTNNNNNNNNNNN 0</td>
</tr>
<tr>
<td>8</td>
<td>300</td>
<td>0.0003</td>
<td>1158 0.949 0.926 0.934 NNNNNNNNTTNNNNNNNNNN 0</td>
</tr>
<tr>
<td>9</td>
<td>c0</td>
<td>0.0003</td>
<td>1155 0.944 0.917 0.926 NNNNNNTTNNNNNNNNNNN 0</td>
</tr>
</tbody>
</table>
### Short Loop, pat 124, br 137681, 0.8908 0.9055 0.7441 (0.03700)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Patterns</th>
<th>Branches</th>
<th>gshre</th>
<th>local</th>
<th>corr</th>
<th>Local History</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5555</td>
<td>0.0040</td>
<td>14753</td>
<td>0.987</td>
<td>0.981</td>
<td>0.912 TNTNTNTNTNTNTNT</td>
</tr>
<tr>
<td>1</td>
<td>aaaa</td>
<td>0.0040</td>
<td>14730</td>
<td>0.859</td>
<td>0.978</td>
<td>0.461 NTTNTNTNTNTNTNT</td>
</tr>
<tr>
<td>2</td>
<td>9249</td>
<td>0.0022</td>
<td>8062</td>
<td>0.997</td>
<td>0.992</td>
<td>0.988 TNNTNNTNNTNNTNT</td>
</tr>
<tr>
<td>3</td>
<td>4924</td>
<td>0.0022</td>
<td>8055</td>
<td>0.997</td>
<td>0.998</td>
<td>0.998 NNTNNTNNTNNTNNT</td>
</tr>
<tr>
<td>4</td>
<td>2492</td>
<td>0.0022</td>
<td>8047</td>
<td>0.993</td>
<td>0.991</td>
<td>0.009 NTTNTNNTNNTNNTN</td>
</tr>
<tr>
<td>5</td>
<td>db6d</td>
<td>0.0013</td>
<td>4864</td>
<td>0.713</td>
<td>0.915</td>
<td>0.065 TNNNTNNTNNTNNT</td>
</tr>
<tr>
<td>6</td>
<td>b6db</td>
<td>0.0013</td>
<td>4713</td>
<td>0.862</td>
<td>0.903</td>
<td>0.926 TNNNTNNTNNTNNT</td>
</tr>
<tr>
<td>7</td>
<td>6db6</td>
<td>0.0012</td>
<td>4640</td>
<td>0.991</td>
<td>0.978</td>
<td>0.970 NTTNTNNTNNTNNT</td>
</tr>
<tr>
<td>8</td>
<td>bbbb</td>
<td>0.0008</td>
<td>3061</td>
<td>0.896</td>
<td>0.936</td>
<td>0.949 TNNNTNNTNNTNNT</td>
</tr>
</tbody>
</table>

### Long Loop?, pat 32, br 185795, 0.9170 0.9052 0.9096 (0.04993)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Patterns</th>
<th>Branches</th>
<th>gshre</th>
<th>local</th>
<th>corr</th>
<th>Local History</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>fffe</td>
<td>0.0025</td>
<td>9204</td>
<td>0.902</td>
<td>0.930</td>
<td>0.913 NTTTTTTTTTTTTTT</td>
</tr>
<tr>
<td>1</td>
<td>8000</td>
<td>0.0025</td>
<td>9198</td>
<td>0.654</td>
<td>0.700</td>
<td>0.705 NNNNNNNNNNNNNNT</td>
</tr>
<tr>
<td>2</td>
<td>7fff</td>
<td>0.0022</td>
<td>8052</td>
<td>0.890</td>
<td>0.817</td>
<td>0.818 TTTTTTTTTTTTTTN</td>
</tr>
<tr>
<td>3</td>
<td>ffbf</td>
<td>0.0018</td>
<td>6800</td>
<td>0.933</td>
<td>0.908</td>
<td>0.920 TTTTTTTTTTTTTTN</td>
</tr>
<tr>
<td>4</td>
<td>feff</td>
<td>0.0018</td>
<td>6782</td>
<td>0.946</td>
<td>0.938</td>
<td>0.942 TTTTTTTTTTTTTTN</td>
</tr>
<tr>
<td>5</td>
<td>ff7f</td>
<td>0.0018</td>
<td>6778</td>
<td>0.949</td>
<td>0.946</td>
<td>0.950 TTTTTTTTTTTTTTN</td>
</tr>
<tr>
<td>6</td>
<td>fdff</td>
<td>0.0018</td>
<td>6738</td>
<td>0.947</td>
<td>0.941</td>
<td>0.946 TTTTTTTTTTTTTTN</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.0018</td>
<td>6690</td>
<td>0.955</td>
<td>0.945</td>
<td>0.942 TNNNNNNNNNNNNNN</td>
</tr>
<tr>
<td>8</td>
<td>fffd</td>
<td>0.0018</td>
<td>6667</td>
<td>0.968</td>
<td>0.966</td>
<td>0.967 TTTTTTTTTTTTTTN</td>
</tr>
</tbody>
</table>
Phase Change, pat 26, br 48190, 0.8453 0.9040 0.8470 (0.01295)

<table>
<thead>
<tr>
<th>% Patterns</th>
<th># Branches</th>
<th>gshre</th>
<th>local corr</th>
<th>Local History</th>
</tr>
</thead>
<tbody>
<tr>
<td>0: c000</td>
<td>0.0012</td>
<td>4554</td>
<td>0.653</td>
<td>0.777 0.680</td>
</tr>
<tr>
<td>1: e000</td>
<td>0.0009</td>
<td>3420</td>
<td>0.714</td>
<td>0.859 0.758</td>
</tr>
<tr>
<td>2: f000</td>
<td>0.0008</td>
<td>2942</td>
<td>0.756</td>
<td>0.888 0.788</td>
</tr>
<tr>
<td>3: fffc</td>
<td>0.0008</td>
<td>2878</td>
<td>0.908</td>
<td>0.960 0.959</td>
</tr>
<tr>
<td>4: f800</td>
<td>0.0007</td>
<td>2642</td>
<td>0.786</td>
<td>0.917 0.827</td>
</tr>
<tr>
<td>5: 3</td>
<td>0.0007</td>
<td>2572</td>
<td>0.968</td>
<td>0.952 0.951</td>
</tr>
<tr>
<td>6: fc00</td>
<td>0.0007</td>
<td>2435</td>
<td>0.815</td>
<td>0.933 0.854</td>
</tr>
<tr>
<td>7: fe00</td>
<td>0.0006</td>
<td>2225</td>
<td>0.836</td>
<td>0.936 0.876</td>
</tr>
<tr>
<td>8: ff00</td>
<td>0.0006</td>
<td>2140</td>
<td>0.856</td>
<td>0.947 0.931</td>
</tr>
<tr>
<td>9: ff80</td>
<td>0.0006</td>
<td>2061</td>
<td>0.854</td>
<td>0.941 0.934</td>
</tr>
</tbody>
</table>

One Way, pat 2, br 2617433, 0.9917 0.9934 0.9897 (0.70337)

<table>
<thead>
<tr>
<th>% Patterns</th>
<th># Branches</th>
<th>gshre</th>
<th>local corr</th>
<th>Local History</th>
</tr>
</thead>
<tbody>
<tr>
<td>0: ffff</td>
<td>0.5151</td>
<td>1916950</td>
<td>0.993 0.996 0.993</td>
<td>TTTTTTTTTTTTTTTTTTTTTTTT 4</td>
</tr>
<tr>
<td>1: 0</td>
<td>0.1882</td>
<td>700483</td>
<td>0.988 0.986 0.982</td>
<td>NNNNNNNNNNNNNNNNNN 4</td>
</tr>
</tbody>
</table>
Two-Level Correlating Predictors

Idea: Base branch decision on . . .
. . . the address of the branch instruction (as in the one-level scheme) . . .
. . . and the most recent branch outcomes.

*History:*
The outcome (taken or not taken) of the most recent branches. Usually stored as a bit vector with 1 indicating taken.

*Pattern History Table (PHT):*
Memory for 2-bit counters, indexed (addressed) by some combination of history and the branch instruction address.
Some Types of Two-Level Predictors

*Global*, a.k.a. *GAg*.

History is global (same for all branches), stored in a global history register (*GHR*).

PHT indexed using history only.

*gshare*

History is global (same for all branches), stored in a global history register (*GHR*).

PHT indexed using history exclusive-ored with branch address.

*gselect*

History is global (same for all branches), stored in a global history register (*GHR*).

PHT indexed using history concatenated with branch address.
Local, a.k.a., PAg.

History is local, BHT stores history for each branch.

PHT indexed using history only.
Global History Example

! Loop always iterates 4 times.
! Branch below never taken.
    bneq r2, SKIP           N
    add f0, f0, f2
SKIP:
    addi r1, r0, #4
LOOP:
    multd f0, f0, f2
    subi r1, r1, #1
    bneq r1, LOOP           T  T  T  N  ...  T  T  T  N  ...
! Cycle                               10  20  30  40  50  110  120  130  140  150
!
! Global History (m=4), X: depends on earlier branches.
! 10  XXXN  Human would predict taken.
! 20  XXNT  Human would predict taken.
! 30  XNTT  Human would predict taken.
! 40  NTTT  Human would predict not taken.
! 50  TTTN
Steps in BHT Use on a Dynamically Scheduled Proc.

Register $r1$ not available until cycle ten$^1$.

Cycle 1: When branch in ID, read BHT and make prediction.

Cycle 1: (Optional) Backup (checkpoint) register map (if present).

Cycle 10: Execute branch in usual way and check prediction.

Cycle 11: If prediction correct, update BHT when branch commits.

Cycle 11: If pred. wrong, start recovery process (does not occur here).

! Predict not taken, not taken.

Cycle: 0 1 2 3 10 11 12 13
bneq r1, TARGET IF ID Q ... B WC
xor r2, r3, r4 IF ID Q EX WB C

... 
TARGET:
and r5, r6, r7

---

$^1$ Perhaps due to a cache miss, or maybe it depended on a long-latency floating-point operation, the reason is not important.
BHT use when branch taken, correctly predicted.

Register \( r1 \) not available until cycle 10.

Cycle 1: When branch in ID, compute target, read BHT and make prediction.

Cycle 10: Execute branch in usual way.

Cycle 11: Check outcome. Correctly predicted.

Cycle 23: Commit branch after \( \text{div} \).

\[
\begin{array}{cccccccccccccc}
\text{Predict taken, taken.} \\
\text{Cycle:} & 0 & 1 & 2 & 3 & 10 & 11 & \ldots & 21 & 22 & 23 \\
\text{div f0,f2, f4} & \text{ID} & \text{Q} & \text{DIV} & \text{DIV WC} \\
\text{bneq r1, TARGET} & \text{IF} & \text{ID} & \text{Q} & \ldots & \text{B} & \text{WB} & \text{C} \\
\text{xor r2, r3, r4} & \text{IFx} \\
\cdots \\
\text{TARGET:} \\
\text{and r5, r6, r7} & \text{IF} & \ldots & \text{C}
\end{array}
\]
BHT use when branch taken, incorrectly predicted, register map not backed up.

Register \texttt{r1} not available until cycle 10.

Cycle 1: When branch in ID, compute target, read BHT and make prediction.

Cycle 10: Compute branch condition.

Cycle 11: Misprediction “discovered.” Because register map not backed up, recovery must wait until commit.

Cycle 23: Start recovery: Squash instructions in reorder buffer, start fetching correct path.

\begin{verbatim}
! Predict not taken, taken. Register map not backed up.
Cycle: 0 1 2 3 10 11 ... 21 22 23
div f0,f2, f4 ID Q DIV 10 11 ... DIV WC
bneq r1, TARGET IF ID Q ... B WB C
xor r2, r3, r4 IF ID Q EX ...

... TARGET:
and r5, r6, r7 IF ....
\end{verbatim}
BHT use when branch taken, incorrectly predicted, register map backed up.

Register \( r1 \) not available until cycle 10.

Cycle 1: When branch in ID, backup (checkpoint) register map, compute target, read BHT and make prediction.

Cycle 10: Compute branch condition.

Cycle 11: Misprediction discovered. Squash reorder buffer past branch, switch to backed up register map, start fetching correct path.

Cycle 23: Branch commits.

---

! Predict not taken, taken. Register map backed up.

<table>
<thead>
<tr>
<th>Cycle:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>10</th>
<th>11</th>
<th>...</th>
<th>21</th>
<th>22</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>div f0, f2, f4</td>
<td>ID</td>
<td>Q</td>
<td>DIV</td>
<td>DIV</td>
<td>WC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bneq r1, TARGET</td>
<td>IF</td>
<td>ID</td>
<td>Q</td>
<td>...</td>
<td>B</td>
<td>WB</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xor r2, r3, r4</td>
<td>IF</td>
<td>ID</td>
<td>Q</td>
<td>EX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

... 

TARGET:

and r5, r6, r7

---
Global History and Dynamic Scheduling

Global history must be accurate.

Why that’s a problem:

! First branch: Predict not taken, taken. Register map backed up.
Cycle: 0 1 2 3 10 11 12 13 ... 21 22 23
div f0, f2, f4 ID Q DIV DIV WC
bneq r1, TARGET IF ID Q ... B WB C
beqz r2, SKIP IF ID Q B ...
xor r2, r3, r4 IF ID EX ...

... TARGET:
and r5, r6, r7 IF ID Q EX ...
beqz r4, LINE1 IF ID Q ...
Cycle: 0 1 2 3 10 11 12 13 ... 21 22 23

Cycle 2: beqz should see global history with bneq not taken.

Global history includes assumption that bneq not taken.
First branch: Predict not taken, taken. Register map backed up.

Cycle: 0 1 2 3 10 11 12 13 ... 21 22 23

```
div f0, f2, f4  ID Q  DIV
bneq r1, TARGET IF ID Q  ... B  WB
beqz r2, SKIP   IF ID Q  B  ...
 xor r2, r3, r4 IF ID Q  EX...
```

TARGET:
```
and r5, r6, r7  IF ID Q  ...
beqz r4, LINE1 IF ID ...
```

Cycle 3: Now global history includes assumption that bneq and first beqz not taken.

Cycle 11: Ooops, bneq misprediction discovered.

Global history has two incorrect assumptions ...
... unless they’re fixed prediction for second beqz won’t be accurate.

Cycle 12: beqz should see global history with bneq taken.
Global History and Dynamic Execution

Global History in Two-Level Predictor with Dynamic Execution

Global history backed up (checkpointed) at each branch.

Predicted outcome shifted into global history.

If misprediction discovered, global history restored from backup . . .

. . . just as the register map can be.
Target Prediction and Folding

**Target Prediction:**
Predicting the outcome and target of a branch.

**Branch Target Buffer:**
A table indexed by branch address holding a predicted target address.

Target Prediction

- Put BTB in IF stage.
- Use PC to read an entry from BTB.
- If valid entry found, replace PC with predicted target.
- With target correctly predicted, zero branch delay.
**Target Prediction Example**

Static scheduled system (for clarity).

<table>
<thead>
<tr>
<th>Cycle:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>bneq r1, TARGET</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
</tr>
<tr>
<td>xor r2, r3, r4</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
</tr>
</tbody>
</table>

**TARGET:**

<table>
<thead>
<tr>
<th>and r5, r6, r7</th>
<th>IF</th>
<th>ID</th>
<th>EX</th>
<th>MEM</th>
<th>WB</th>
</tr>
</thead>
</table>

Cycle 0

BTB lookup and prediction. Predict taken.

Target from BTB will be clocked into PC.
Target Prediction Example, continued.

Static scheduled system (for clarity).

<table>
<thead>
<tr>
<th>Cycle:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>bneq r1, TARGET</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
</tr>
<tr>
<td>xor r2, r3, r4</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td></td>
<td></td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TARGET:

<table>
<thead>
<tr>
<th>and r5, r6, r7</th>
<th>IF</th>
<th>ID</th>
<th>EX</th>
<th>MEM</th>
<th>WB</th>
<th>IF</th>
<th>X</th>
</tr>
</thead>
</table>

Cycle 1

Start fetching predicted target.

Execute branch instruction (in ID).

Check predicted outcome and predicted target.

Correct predictions, continue execution.
Target Prediction Example, continued.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>bneq</td>
<td>r1</td>
<td>TARGET</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td>MEM</td>
<td>WB</td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
</tr>
<tr>
<td>xor</td>
<td>r2</td>
<td>r3</td>
<td>r4</td>
<td></td>
<td></td>
<td></td>
<td>IF</td>
<td>ID</td>
<td>EX</td>
<td></td>
</tr>
</tbody>
</table>

TARGET:

| and   | r5 | r6 | r7 | IF | ID | EX | MEM | WB | IF | X  |

Cycle 10

BTB lookup and prediction. Predict taken.

Target from BTB will be clocked into PC.

Cycle 11

Start fetching predicted target.

Execute branch instruction (in ID).

Ooops, incorrect outcome prediction . . .

... replace target with \texttt{nop} . . .

... and clock correct target into PC.
Target Prediction for Register-Indirect CTI

What BTB predicts for branch instructions:

That instruction will be a CTI.

If CTI is a branch, that branch is taken.

CTI target.

For branches and non-indirect jumps (j, jal)... 
... predicting target is easy, since target always same.

\[
\text{bneq r1, LOOP} \quad \text{! Target always PC + 4 + 4 * LOOP}
\]

\[
\text{j LINEJ} \quad \text{! Target always PC + 4 + 4 * LINEJ}
\]

For register-indirect jumps (jr, jalr)...
... prediction depends on predictable behavior.

\[
\text{jr r1} \quad \text{! Target is in r1. Can be different each time.}
\]

\[
\text{jalr r1} \quad \text{! Target is in r1. Can be different each time.}
\]
Behavior of Register-Indirect Jumps

Predictability depends on how jumps used.

Major Uses

- Procedure Passed as Parameter

  For example, function passed to the C library’s `qsort`.

  These rarely change so target is predictable.

- Case Statements

  These change, and so prediction more difficult.
Indirect Jump Target Prediction

Separate techniques used for procedure returns and other indirect jumps.

Return Address Prediction

Keep a stack of (what appear to be) return addresses.

Other Indirect Jumps Prediction

Predict last target.

Use global branch history to index BTB.
Predict Return Address

Used for return instruction. (An instruction used for a procedure return, which may not have the mnemonic `return`).

Operation

    Hardware keeps a stack of return addresses.

    BTB stores whether instruction is a return.

    When a call instruction encountered push return address on stack.

    When BTB identifies instruction as a return target address is popped off stack.

Effectiveness

    Works fairly well.

    Can be confused when returns skipped (as with long jumps).

    Costly to implement precisely with dynamic scheduling.
Predict Last Target

Can be used for everything except return instructions.

Last time instruction executed target address stored in BTB.

If entry found and predicted taken (for a branch), last target address used.

Effectiveness:

Perfect for non-indirect jumps and branches (if taken).

Reasonably effective on indirect branches.

Use Global History

Can be used for everything except return instructions.

Much more effective on than last target.
Target Prediction Example

Consider code for C `switch` statement:

```
! Possible code for a switch statement.
! switch( r2 ) { case 0: foo(); break; case 1: bar(); break; ... }
! Set r1 to base of switch address table.
  lhi  r1, #0x1234
  ori  r1, r1, #0x5670
! Multiply switch index by stride of table (4 bytes per address).
  slli  r3, r2, #2
! Get address of case code address.
  add   r1, r1, r3
! Get case code address.
  lw    r4, 0(r1)
! Jump to case code.
  jr     r4
```

If r2 rarely changes, jr predictable.
Possible BTB Contents

Target address.

History information (replaces BHT).

Tag, to detect collisions.