LAB 3: BRANCH PREDICTORS

- In this lab students are expected to understand the working of the branch prediction unit in the simplescalar tool.
- The benchmark is the same, compress.ss
- The objective of this lab, as before, is understanding and optimization.
- As will be discussed in class, there are a few types of branch predictors
  - Always Taken – Branch is always taken
  - Always Not taken – Branch is always not taken
  - Bimodal - some lower bits of the PC are used as addresses for 2 bit saturating counters whose MSB is used to predict if the branch is taken or not
  - 2 level BP or correlating BP: explained in next slide
  - Combined (not in this Lab)
Different types of 2 level Branch Predictors are shown in figure.

There are 2 tables, the one on the left is the branch history table and the one on the right is the pattern history table.

Types
- **GAg**: Branch history is global with Hist Size $W$.
  Pattern history is addressed with $W$ bits.
- **GAp**: Branch history is global with Hist Size $W$.
  Pattern history is addressed with those $W$ bits + extra PC LSB bits. Adds semblance of Locality.
- **PAg**: Separate Branch History for $N$ addresses. $\log_2 N$ bits of PC taken.
  Branch history of each address, $W$ bits.
  Only uses Branch history to address Pattern History.
- **PAp**: Separate Branch History for $N$ addresses. $\log_2 N$ bits of PC taken.
  Branch history of each address, $W$ bits.
  Combines LSBs of PC and Branch History to address Pattern History.

![Figure 6. 2-level adaptive predictor structure](image)

<table>
<thead>
<tr>
<th>predictor</th>
<th>H1_size</th>
<th>hist_size</th>
<th>H2_size</th>
<th>xor</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAg</td>
<td>1</td>
<td>$W$</td>
<td>$2^W$</td>
<td>0</td>
</tr>
<tr>
<td>GAp</td>
<td>1</td>
<td>$W$</td>
<td>$&gt;2^W$</td>
<td>0</td>
</tr>
<tr>
<td>PAg</td>
<td>$N$</td>
<td>$W$</td>
<td>$2^W$</td>
<td>0</td>
</tr>
<tr>
<td>PAp</td>
<td>$N$</td>
<td>$W$</td>
<td>$N*2^W$</td>
<td>0</td>
</tr>
<tr>
<td>gshare</td>
<td>1</td>
<td>$W$</td>
<td>$2^W$</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Branch predictor parameters
OBJECTIVE

- Plot curves for mis-prediction rate by varying the different parameters
  - Bimodal: Vary size of the Pattern history Table
  - 2 Level
    - Calculate area of branch predictor in term of bits
      - Assume entries in pattern history tables are always 2 bits
    - Plot mis-prediction rate by varying the parameters in the 2 level predictor
  - Used total area numbers (in bits) to be the following
    - $2^8$ to $2^{16}$ in logarithmic steps
    - For each fixed area, figure out optimal points for minimizing mis-prediction rate by varying one of the 2 parameters (N,W). (since area is fixed, varying N will vary W as system is constrained)
  - Using sim-profile with “–all” argument, look at the looping patterns and explain all of the results of the above experiments as well as the curves that you get
  - For each of the total area numbers as mentioned above($2^8$ to $2^{16}$), explain why different predictors have better/worse mis-prediction rates
GRADING

- Again, grading will be done based on explanation and reasoning and not the answers.
- Please do not submit numbers without reasons. The objective of this is to make you think about branch prediction.
- As usual, this week’s bonus question:
  - Perform energy estimation of the branch predictor using the Wattch tool.
  - Get power numbers from watch.
  - Multiply by number of cycles and assume CCT to be the same as LAT from the Cache Lab.
  - Bonus points will be awarded to students who not just show the data, but EXPLAIN WHY they got these results.
- Lab 3 is due 1.5 weeks from now, Friday 21\(^{th}\) of February via EMAIL PDF reports.