ECE552 PRA0102 - Station 8
Ting-Hao (Tim) Cheng (997289090)
Isaiah Tan (999004937)

Microbench

Compiled without additional flags. We created the microbenchmark to branch unless it's on the 4th (i%3) or 6th (i%5) iteration.

```
for (i = 0; i < n; i++) {
        if (i%3 == 0)
                a ++;
        if (i%5 == 0)
                b ++;
}
```

for n = 100; MPKI: 2-bit saturating = 14.855, 2-level = 16.261
for n = 100000; MPKI: 2-bit saturating = 15.871, 2-level = 0.491

In the 100 iterations run, it makes sense that the 2-level results will be similar to the 2-bit saturating MPKI (including the branches outside the for loop scope). the 2-level require multiple setup iterations before the history of the for loop is stored.Compared to the 100000 iteration run, we can see that the 2-level is near perfect, which is expected as once the prediction tables have correctly identified the the direction based on history, the predictions should have no inaccuracy

MPKI over 8 benchmarks

Benchmark	**2-bit saturating**	**2-level**	**open-ended**
astar	24.639	11.903	5.073
bwaves	7.886	7.146	5.110
bzip2	8.166	8.651	7.734
gcc	21.079	14.824	5.099
gromacs	9.088	7.484	5.182
hmmer	13.567	14.872	12.558
mcf	24.387	13.494	10.353
soplex	7.107	6.819	5.154

Open-ended Predictor

Our idea is a tournament/hybrid predictor. We used a 2-bit saturating chooser w/4096 entries, indexed by a GHR. The chooser decides between a PAp and gshare predictor. The PAp predictor has 2048 history entries, 8 prediction tables, 2048 entries per prediction table, and uses 3-bit saturating counters. The PC is used to select the prediction table, and the history entry. The history value retrieve is used to index the selected prediction table. The gshare predictor uses the GHR (from the chooser) and has 4096 entries of 2-bit saturating counters. The gshare table is indexed by XORing the GHR with the PC's least significant bits. As branches could yield different results (even on the same PC) with different incoming global history, gshare accounts for this by the XOR, to use different predictions based on global history.

Size: 4096*2 (gshare)+4096*2 (chooser)+12 (global history)+2048*11 (2-level history table)+8*2048*3 (2-level prediction table) = 88076 bits = 86.01 Kb

CACTI Results

Two level Predictor

Note that there are 8 predictor tables, indexed by 3bits of PC. The measurements are for each table.

	Config	Area	Access Latency	Leakage Power
History Table (2level 1 cfg)	Type: ram, Size: 512, Block Size: 1, I/O Width: 8, Tag: -	0.0391054 x 0.0269215	0.163585	0.195006
Predictor Table (2level 2 cfg)	Type: ram, Size: 64, Block Size: 1, I/O Width: 8, Tag: N/A (per predictor table)	0.0256629 x 0.00746754 (Per predictor table)	0.117805	0.0299666 (Per predictor table)

Open-ended Predictor

Note that there are 8 predictor tables, indexed by 3 bits of PC. The measurements are for each table.

	Config (per structure)	Area (mm)	Access Latency	Leakage Power (mW)
Chooser & Gshare predictor table (open 1 cfg)	Type: ram, Size: 4096, Block Size: 1, I/O Width: 8, Tag: N/A	0.0391054 x 0.0269215	0.163585	0.195006
2level History Table (open 2 cfg)	Type: ram, Size: 4096, Block Size: 2, I/O Width: 16, Tag: N/A	0.135196 x 0.050676	0.206818	1.61593
2level Predictor Table (open 3 cfg)	Type: ram, Size: 2048, Block Size: 1, I/O Width: 8, Tag: N/A (Per predictor table)	0.0700302 x 0.0515418 (Per predictor table)	0.206091	0.833957 (Per predictor table)

Work Delegation

Isaiah: implemented 2-level local history predictor, designed open-ended, implemented open-ended
Tim: implemented 2-bit saturating predictor, designed open-ended, created microbenchmark and CACTI config files